ARO 36224.1-PH-CF

Volume 65
Number 6, 1997

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

International Journal of QUANTUM CHEMISTRY

Editor-in-Chief
PER-OLOV LÖWDIN

Editors
ERKKI BRÄNDAS
YNGVE ÖHRN

Associate Editors
OSVALDO GOSCINSKI
STEN LUNELL
JOHN R. SABIN
MICHAEL C. ZERNER

Proceedings of the
International Symposium on the
Application of Fundamental Theory to

Problems of Biology and Pharmacology

Held at the Ponce de Leon Resort,
St. Augustine, Florida, March 1–7, 1997

Editor-in-Chief: Per-Olov Löwdin
Special Editors: Yngve Öhrn
John R. Sabin
Michael C. Zerner

A Wiley-Interscience Publication
John Wiley & Sons, Inc.

ISSN 0020-7608

International Journal of

Quantum Chemistry

Editor-in-Chief

Per-Olov Löwdin
University of Florida at Gainesville, USA
Uppsala University, Sweden

Editors

Erkki Brändas
Uppsala University, Sweden
Yngve Öhrn
University of Florida at Gainesville, USA

Associate Editors

Osvaldo Goscinski
Uppsala University, Sweden
Sten Lunell
Uppsala University, Sweden
John R. Sabin
University of Florida at Gainesville, USA
Michael C. Zerner
University of Florida at Gainesville, USA

Honorary Editors

Gerhard Herzberg
National Research Council, Ottawa, Ontario, Canada
Kenichi Fukui
Institute for Fundamental Chemistry at Kyoto, Japan
Jerome Karle
Naval Research Laboratory at Washington, DC, USA
Rudy Marcus
California Institute of Technology at Pasadena, USA

Editorial Board

Jiri Čížek
University of Waterloo, Ontario, Canada
Enrico Clementi
Université Louis Pasteur, Strasbourg, France
Raymond Daudel
Académie Européene de Arts, des Sciences et des Lettres, Paris, France
Ernest Davidson
Indiana University at Bloomington, USA
George G. Hall
University of Nottingham, UK
Laurens Jansen
Kusnacht, Switzerland
Norman H. March
University of Oxford, UK
Roy McWeeny
Università di Pisa, Italy
Saburo Nagakura
Graduate University for Advanced Studies, Yokohama, Japan
Kimio Ohno
Hokkaido Information University, Japan
Josef Paldus
University of Waterloo, Ontario, Canada
Robert G. Parr
University of North Carolina at Chapel Hill, USA
Ruben Pauncz
Technion, Haifa, Israel
John A. Pople
Northwestern University at Evanston, Illinois, USA
Alberte Pullman
Institut de Biologie Physico-Chimique, Paris, France
Paul von Ragué Schleyer
Universität Erlangen-Nürnberg, Erlangen, Germany
Harrison Shull
Naval Postgraduate School, Monterey, California, USA
Tang Au-Chin
Jilin University, Changchun, China
Rudolf Zahradník
Czech Academy of Sciences, Prague, Czech Republic

Advisory Editorial Board

Teijo Åberg
Helsinki University of Technology, Espoo, Finland
Michail V. Basilevsky
Karpov Institute of Physical Chemistry, Moscow, Russia
Axel D. Becke
Queen's University, Kingston, Ontario, Canada
Gian Luigi Bendazzoli
Università di Bologna, Italy
Geza Biczo
Hungarian Academy of Sciences, Budapest, Hungary
Jerzy Cioslowski
The Florida State University at Tallahassee, USA
Timothy Clark
Universität Erlangen-Nürnberg, Germany
Giorgina Corongiu
Centro di Richerche, Sviluppo e Studi Superiori in Sardegna, Cagliari, Italy
Mireille Defranceschi
DPEL/SERGD/LMVT, Fontenay Aux Roses, France
Karl F. Freed
The University of Chicago, Illinois, USA
Peter Fulde
Max-Planck-Institut für Physik Komplexer Systeme, Dresden, Germany
Odd Gropen
University of Tromsø, Norway
Trygve Helgaker
University of Oslo, Norway
Ming-Bao Huang
Academia Sinica, Beijing, People's Republic of China
James T. Hynes
University of Colorado at Boulder, USA
Mu Shik Jhon
Korea Advanced Institute of Science and Technology, Seoul, Korea
Hiroshi Kashiwagi
Kyushu Institute of Technology, Fukuoka, Japan
Peter A. Kollman
University of California at San Francisco, USA
Eugene S. Kryachko
Academy of Sciences of Ukraine,
Sven Larsson
Chalmers University of Technolog Gothenburg, Sweden
Lucas Lathouwers
Universitair Centrum (RUCA), Ant
Shyi-Long Lee
National Chung Chang University, Taiwan, Republic of China
Claude Leforestier
Université Paris-Sud, Orsay, Franc
Josef Michl
University of Colorado at Boulder,
Nimrod Moiseyev
Israel Institute of Technology, Isra
John D. Morgan III
University of Delaware at Newark,
Cleanthes A. Nicolaides
National Hellenic Research Founc
J. Vincent Ortiz
Kansas State University at Manha
Lars Pettersson
University of Stockholm, Sweden
Leon Phillips
University of Canterbury, Christchurch, New Zealand
Martin Quack
ETH Zürich, Switzerland
Leo Radom
Australian National University, Au
William Reinhardt
University of Washington at Seatt
Sten Rettrup
H. C. Ørsted Institut, Copenhagen
C. Magnus L. Rittby
Texas Christian University at Fort
Michael Robb
King's College, London, UK
Mary Beth Ruskai
University of Massachusetts at Lc
Harold Scheraga
Cornell University at Ithaca, New
Vipin Srivastava
University of Hyderabad, India
Nicolai F. Stepanov
Moscow State University, Russia
Jiazong Sun
Jilin University, Changchun, People's Republic of China
Colin Thomson
University of St. Andrews, Scotlar
Donald G. Truhlar
University of Minnesota at Minne
Frank Weinhold
University of Wisconsin at Madis
Harel Weinstein
Mount Sinai School of Medicine, New York, USA
Peter Wolynes
University of Illinois at Urbana, U
Robert E. Wyatt
The University of Texas at Austin,

Editorial Production, John Wiley & Sons: Sheila Kaminsky

. Vol. 65, No. 6

International Journal of QUANTUM CHEMISTRY

Quantum Biology Symposium No. 24

Proceedings of the International Symposium on the Application of Fundamental Theory to Problems of Biology and Pharmacology

Held at Ponce de Leon Resort, St. Augustine, Florida, March 1–7, 1997

Editor-in-Chief: Per-Olov Löwdin
Special Editors: Yngve Öhrn, John R. Sabin, and Michael C. Zerner

an Interscience® Publication
published by JOHN WILEY & SONS

DTIC QUALITY INSPECTED 2

19980519 056

The *International Journal of Quantum Chemistry* (ISSN 0020-7608) is published semimonthly with one extra issue in January, March, May, July, August, and November by John Wiley & Sons, Inc., 605 Third Avenue, New York, New York 10158.

Copyright © 1997 John Wiley & Sons, Inc. All rights reserved. No part of this publication may be reproduced in any form or by any means, except as permitted under section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the publisher, or authorization through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (508) 750-8400, fax (508) 750-4470. Periodicals postage paid at New York, NY, and at additional mailing offices.

The code and the copyright notice appearing at the bottom of the first page of an article in this journal indicate the copyright owner's consent that copies of the article may be made for personal or internal use, or for the personal or internal use of specific clients, on the condition that the copier pay for copying beyond that permitted by Sections 107 or 108 of the US Copyright Law.

This consent does not extend to the other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective work, or for resale. Such permission requests and other permission inquiries should be addressed to the Permissions Dept.

Subscription price (Volumes 61–65, 1997): \$3,860.00 in the US, \$4,160.00 in Canada and Mexico, \$4,355.00 outside North America. All subscriptions outside US will be sent by air. Personal rate (available only if there is an institutional subscription): \$190.00 in North America, \$370.00 outside North America. Subscriptions at the personal rate are available only to individuals. Payment must be made in US dollars drawn on a US bank. Claims for undelivered copies will be accepted only after the following issue has been received. Please enclose a copy of the mailing label. Missing copies will be supplied when losses have been sustained in transit and where reserve stock permits.

Please allow four weeks for processing a change of address. For subscription inquiries, please call (212) 850-6645; e-mail: SUBINFO@Wiley.com.

Postmaster: Send address changes to *International Journal of Quantum Chemistry*, Susan Swanson, Director, Subscription Fulfillment and Distribution, Subscription Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158.

Advertising Sales: Inquiries concerning advertising should be forwarded to Susan Levey, Advertising Sales, John Wiley & Sons, Inc., 605 Third Avenue. New York, NY 10158; (212) 850-8832. Advertising Sales, European Contact: Michael Levermore, Advertising Manager, John Wiley & Sons, Ltd., Baffins Lane, Chichester, Sussex PO 19 1UD, England.

Reprints: Reprint sales and Inquiries should be directed to the customer service department, John Wiley & Sons, Inc. 605 Third Ave., New York, NY 10158. Tel: 212-850-8776.

Manuscripts should be submitted in triplicate and accompanied by an executed Copyright Transfer Form to the Editorial Office, *International Journal of Quantum Chemistry*, Quantum Chemistry Group, Uppsala University, Box 518, S-75120, Uppsala, Sweden. Authors may also submit manuscripts to the Editorial Office, *International Journal of Quantum Chemistry*, Quantum Theory Project, Williamson Hall, University of Florida, Gainesville, Florida 32611. **Information for Contributors** appears in the first and last issue of each volume. **All other correspondence** should be addressed to the *International Journal of Quantum Chemistry*, Publisher, Interscience Division, Professional, Reference, and Trade Group, John Wiley & Sons, Inc., 605 Third Avenue, New York, NewYork 10158, U.S.A. The contents of this journal are indexed or abstracted in *Chemical Abstracts, Chemical Titles, Chemical Database, Current Contents/Physical, Chemical, and Earth Sciences, Research Alert (ISI), Science Citation Index (ISI),* and *SCISEARCH Database (ISI).*

This paper meets the requirements of ANSI/NISO Z39. 48–1992 (Permanence of Paper). ♾

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	February 1998	Final Report 28 Jan 97 - 27 Jan 98

4. TITLE AND SUBTITLE

Support of the 1997 Sanibel Symposium Proceedings of the International Symposium on the Application of Fundamental Theory to Problems of Biology and Pharmacology

5. FUNDING NUMBERS

DAAG55-97-1-0020

6. AUTHOR(S)

Yngve Ohrn (principal investigator)

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)

University of Florida
Gainesville, FL 32611

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

ARO 36224.1-PH-CF

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

ABSTRACT NOT AVAILABLE

14. SUBJECT TERMS

15. NUMBER IF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

International Journal of

Quantum Chemistry

Volume 65
Number 6, December 20, 1997

Contents

(continued)

Introduction

The 37th Annual Sanibel Symposium, organized by the faculty, students, and staff of the Quantum Theory Project of the University of Florida, was held on March 1–7, 1997. The meeting was again held at the Ponce de Leon Conference Center in St. Augustine, Florida.

The symposium followed the established format with plenary and poster sessions. This year, the schedule was shortened somewhat with a compact seven-day integrated program of quantum biology, quantum chemistry, and condensed matter physics. The topics of the sessions covered by these proceedings include Quantum Biology, Quantum and Classical Molecular Dynamics, Protein Structure and Folding, Monte Carlo Simulations, and Free Energy Calculations of Biological Molecules.

The articles have been subjected to the ordinary refereeing procedures of the *International Journal of Quantum Chemistry*. The articles presented in the sessions on quantum chemistry, condensed matter physics, and associated poster sessions are published in a separate issue of the *International Journal of Quantum Chemistry*.

The organizers acknowledge the following sponsors for their support of the 1997 Sanibel Symposium:

- Army Research Office and U. S. Army Edgewood RD&E Center through Grant #DAAG55-97-1-0020:

  "The views, opinions, and/or findings contained in the report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation."
- The Office of Naval Research through Grant #N00014-97-1-0320:

  "This work relates to Department of the Navy Grant #N00014-97-1-0320 issued by the Office of Naval Research. The United States Government has the royalty-free license throughout the world in all copyrightable material contained herein."
- IBM Corporation.
- Hypercube, Inc.
- Silicon Graphics.
- The University of Florida.

Very special thanks to the staff of the Quantum Theory Project of the University of Florida for handling the numerous administrative, clerical, and practical details. The organizers are proud to recognize the contributions of Mrs. Judy Parker, Ms. Sharon Stellato, Ms. Sandra Weakland, and Mr. Greg Pearl. All the graduate students of the Quantum Theory Project, who served as "gofers," are gratefully recognized for their contributions to the 1997 Sanibel Symposium.

N. Y. Öhrn
J. R. Sabin
M. C. Zerner

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061031-01

Ab Initio and Molecular Mechanics Conformational Analysis of Neutral L-Proline

MICHAEL RAMEK,[1] ANNE-MARIE KELTERER,[1] SONJA NIKOLIĆ[2]

[1] *Institut für Physikalische und Theoretische Chemie, Technische Universität Graz, A-8010 Graz, Austria*

[2] *Ruđer Bošković Institute, P.O.B. 1016, HR-10 000 Zagreb, Croatia*

Received 1 March 1997; revised 21 May 1997; accepted 23 May 1997

ABSTRACT: The energetically low-lying parts of the potential energy surface of L-proline were investigated by ab initio (RHF/6-311++G**) calculations. The results are discussed with respect to the parametrization of the MM3 force field and in comparison with those obtained earlier for glycine and α-alanine, on the one hand, and for *N*-acetyl-L-proline amide, on the other hand. © 1997 John Wiley & Sons, Inc. Int J Quant Chem **65**: 1033–1045, 1997

Introduction

Amino acids have often been the target of quantum chemical structure investigations for a number of reasons. One of these reasons is the fact that amino acids form zwitterions in the solid state and in polar media, whereas the neutral form is more stable for isolated molecules. Most of the routine experimental techniques for structure determination, like X-ray crystallography, offer easy access to the zwitterionic structure. Experimental structure studies of the neutral form require much more sophisticated techniques. Quantum chemical structure investigations are a perfect tool in this situation, since their basic approach is to consider a single, isolated molecule. The case of glycine, in which the structure of the most stable neutral conformer was first predicted on the basis of ab initio calculations [1] and, later, based on the predicted values, confirmed experimentally [2, 3], proved this advantage of quantum chemical methods already almost 20 years ago. Another reason for theoretical structure investigations of amino acids is, of course, the biological importance of amino acids as the building blocks for peptides and proteins and the need to understand and to describe correctly the weak interactions between the functional groups in these molecules as a prerequisite for a computerized molecular modeling of this class of compounds.

High-level quantum chemical ab initio calculations were reported recently for a number of amino acids, especially for glycine, alanine, serine, and cysteine [4–8], which all have simple side chains. For proline, in which the side chain is linked to the

Correspondence to: M. Ramek.

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061033-13

amino nitrogen atom, an early ab initio study was performed a decade ago by Sapse et al. [9]. Most of that work was based on data obtained with the STO-3G basis set and subject to several artificial restrictions on the geometry parameters, which were seemingly necessary to be able to perform the calculations with the computational resources available at that time. As a consequence of these restrictions, the five-membered pyrrolidine ring of proline was predicted to be planar in the free acid as well as in the *N*-formyl and the *N*-acetyl amide.

Recently, the neutral form of proline has become the focus of experimentally dominated interest: Using a matrix isolation technique, the molecular structure was experimentally confirmed to be nonzwitterionic [10], and the vibrational frequencies of proline and hydroxyproline were evaluated with the help of ab initio-optimized geometries [11]. Part of that interest in proline is, of course, due to the unique nature of proline among the naturally occurring amino acids, which is caused by the ring structure with only one hydrogen atom bonded to the nitrogen atom. Since this hydrogen atom is eliminated when proline becomes part of a peptide chain, proline has only limited possibilities to form the hydrogen bonds that stabilize the helices and sheet structures in peptides and proteins. Hence, proline is a delimiter of these structural units, which makes it important for the secondary structure of proteins and peptides [12–16].

We investigated the potential energy surface (PES) of L-proline using various theoretical methods. The results of previous RHF calculations on neutral ω-amino acids [17–19] and ω-hydroxy acids [20–22] showed that conformers with a *syn*-periplanar orientation of the groups C=O and O—H are 35–40 kJ/mol more stable than are those with the groups C=O and O—H in *anti*-periplanar orientation [23]. In view of this, the present study was focused on conformers with a *syn*-periplanar arrangement of the COOH group; only two conformers with an *anti*-periplanar COOH group were included, namely, those that form an intramolecular N···H—O hydrogen bond.

The core of this work is a description of these energetically low-lying parts of the PES in terms of local minima and reaction paths as obtained via RHF [24] calculations with the standard basis set 6-311++G** [25–27]. In addition, the PES was searched for local minima by molecular mechanics calculations using the MM3(94) force field [28]. This is of special interest because the torsion parameter N—C—C—O (type 8-3-1-75), which is important for the description of the position of the carboxyl group relative to the amino group in neutral amino acids, is missing in the MM3(94) parameter set.

Computational Details

At the ab initio level, all stationary geometries were fully optimized to remaining maximum and root mean square (rms) gradients less than 1.0×10^{-4} and 0.33×10^{-4} H/Bohr, respectively. The nature of all stationary geometries was confirmed by calculating the eigenvalues of the Hessian matrix; according to the computational resources available to the authors, the Hessian matrix was computed by double numerical differentiation of analytical first derivatives. The program GAMESS [29] was used for the ab initio calculations on a number of machines.

The MM3(94) force-field program was used on a DEC Alpha 3000-900 workstation. The parameter sets implemented in the program are dated 10-Jan-94. In addition, the hydrogen-bond parameters (type 23-77) were added with the MM3(96) values of $\epsilon = 1.1$ kcal/mol and $r = 2.39$ Å [30]. To find the local minima, the stochastic search routine was used; starting from the located minima, the potential energy curves were traced with the dihedral driver by block diagonal Newton–Raphson minimization. The energy gradients were less than 1×10^{-5} kcal/mol. For the characterization of local minima and transition states, the vibrational spectra were calculated with the full Newton–Raphson method. Structures with "restricted internal rotations" below 20 cm^{-1} were ignored because of the flat nature of the PES in the respective region. The atom labeling used throughout this work is defined in Figure 1.

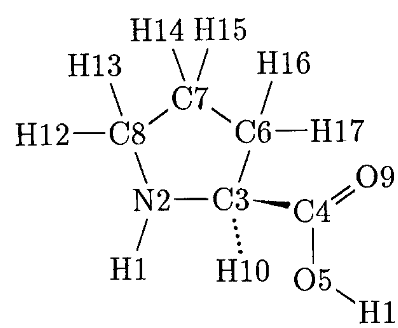


FIGURE 1. Definition of atom labels used in this work.

RHF Results

Ten conformers were located in the PES of neutral L-proline, which are labeled **1**, **2**, ..., **10** according to the energy. The geometries of these conformers are sketched in Figure 2 and the geometry data (bond lengths, valence, and torsion angles) and the energies are collected in Table I.

Due to the restrictions imposed by the pyrrolidine ring, these 10 conformers may be characterized by only four torsion angles: First, the torsion

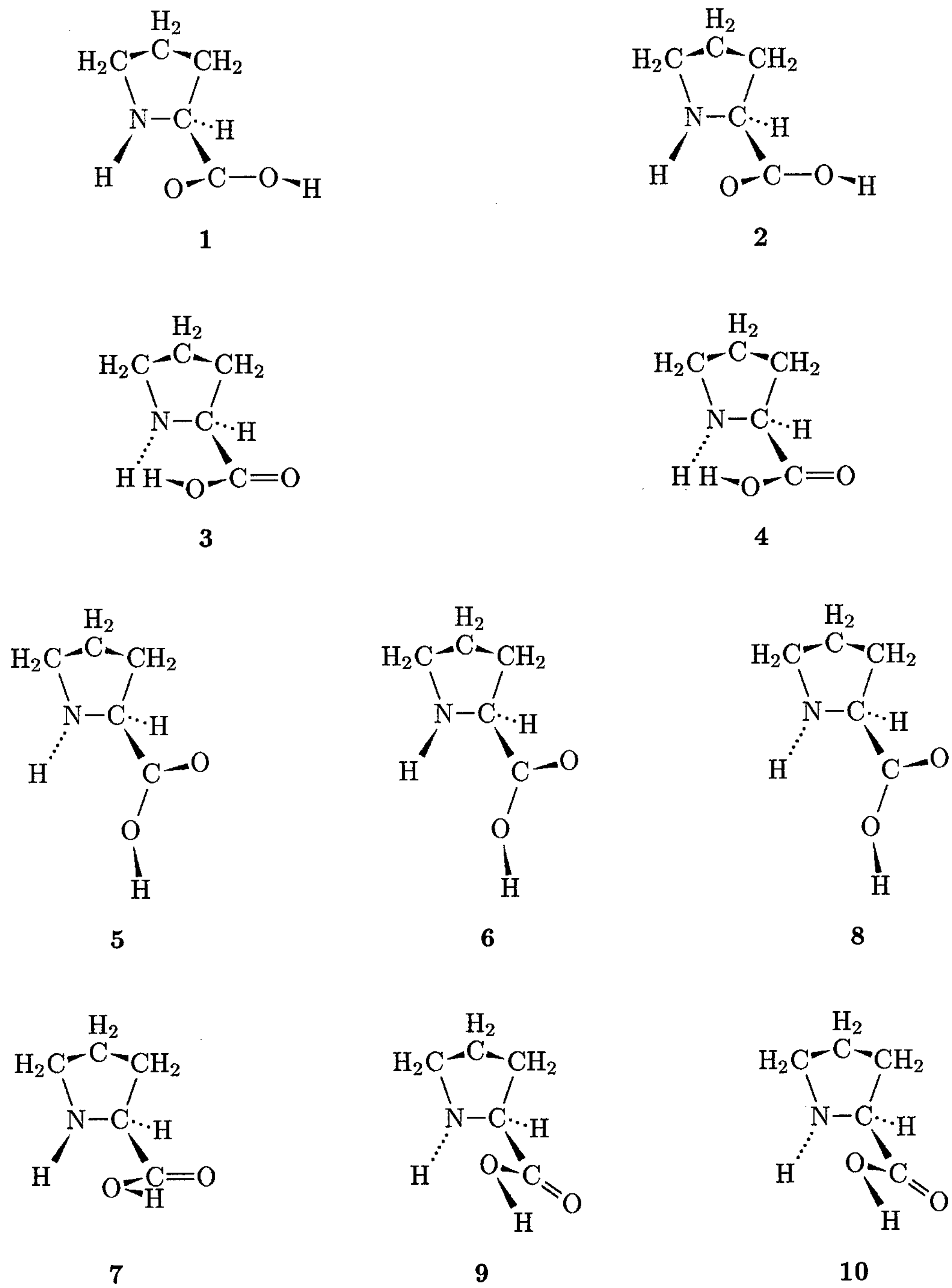


FIGURE 2. Sketches and nomenclature of the L-proline conformers discussed in this work.

TABLE I

Geometry data (Å, degree) and relative energies (kJ / mol) of the local minima in the RHF / 6-311 + +G** PES of L-proline; zero energy corresponds to −398.8817320 au.

	1	2	3	4	5	6	7	8	9	10
E	0.000	2.264	3.104	4.448	6.353	8.348	8.577	8.974	9.197	10.682
H1—N2	0.9976	0.9964	0.9971	0.9968	0.9965	0.9976	0.9964	0.9956	0.9963	0.9963
N2—C3	1.4520	1.4505	1.4658	1.4657	1.4640	1.4484	1.4541	1.4605	1.4615	1.4633
C3—C4	1.5137	1.5191	1.5326	1.5249	1.5185	1.5293	1.5168	1.5147	1.5183	1.5161
C4—O5	1.3286	1.3286	1.3169	1.3182	1.3301	1.3318	1.3306	1.3305	1.3279	1.3276
C3—C6	1.5514	1.5428	1.5420	1.5421	1.5393	1.5413	1.5546	1.5250	1.5443	1.5502
C6—C7	1.5318	1.5295	1.5304	1.5282	1.5333	1.5477	1.5328	1.5295	1.5322	1.5317
C7—C8	1.5238	1.5304	1.5256	1.5288	1.5237	1.5308	1.5235	1.5339	1.5243	1.5287
C4—O9	1.1839	1.1840	1.1811	1.1808	1.1827	1.1828	1.1833	1.1822	1.1845	1.1844
C3—H10	1.0862	1.0857	1.0842	1.0849	1.0877	1.0870	1.0830	1.0909	1.0838	1.0824
O5—H11	0.9461	0.9461	0.9512	0.9507	0.9459	0.9461	0.9463	0.9459	0.9456	0.9456
C8—H12	1.0919	1.0843	1.0837	1.0862	1.0841	1.0938	1.0916	1.0858	1.0842	1.0861
C8—H13	1.0840	1.0868	1.0894	1.0835	1.0928	1.0835	1.0841	1.0856	1.0921	1.0839
C7—H14	1.0857	1.0843	1.0835	1.0874	1.0840	1.0839	1.0854	1.0873	1.0842	1.0876
C7—H15	1.0841	1.0867	1.0867	1.0841	1.0851	1.0838	1.0841	1.0839	1.0839	1.0847
C6—H16	1.0860	1.0823	1.0811	1.0867	1.0805	1.0805	1.0858	1.0839	1.0827	1.0841
C6—H17	1.0822	1.0853	1.0855	1.0814	1.0855	1.0825	1.0820	1.0833	1.0855	1.0833
H1—N2—C3	111.95	111.06	112.60	111.29	113.20	113.13	113.60	112.72	112.72	110.23
N2—C3—C4	113.15	112.35	111.56	112.58	107.76	114.64	116.17	109.53	108.97	110.94
C3—C4—O5	111.71	111.96	115.51	115.31	111.09	112.30	113.91	111.22	113.47	113.24
N2—C3—C6	105.76	105.02	105.52	105.37	105.42	103.98	105.06	104.20	105.23	106.55
C3—C6—C7	103.39	102.68	103.13	102.73	104.21	104.71	103.93	101.98	104.09	103.41
C6—C7—C8	102.32	102.22	102.40	102.12	102.41	104.31	102.47	102.87	102.38	101.97
C3—C4—O9	125.80	125.69	121.81	121.96	126.57	125.76	123.86	126.38	124.24	124.39
N2—C3—H10	110.42	111.57	112.15	111.64	112.15	110.14	110.45	112.02	112.31	110.66
C4—O5—H11	108.94	108.86	107.73	108.02	108.83	108.70	108.53	108.86	108.68	108.69
C7—C8—H12	110.13	112.74	113.20	110.39	113.26	110.41	110.31	110.81	112.98	110.34
C7—C8—H13	113.07	110.41	110.13	113.54	110.18	113.10	113.07	112.24	110.17	113.78
C6—C7—H14	110.52	112.99	113.15	110.23	112.86	110.30	110.56	110.01	112.77	110.48
C6—C7—H15	112.71	110.23	110.57	113.14	110.39	112.47	112.58	113.17	110.37	112.77
C3—C6—H16	110.38	112.88	112.21	110.21	112.01	110.82	110.41	109.59	112.94	110.43
C3—C6—H17	112.00	108.66	109.24	111.74	109.11	109.94	111.59	112.46	108.71	111.61
H1—N2—C3—C4	−16.57	−16.47	−110.59	−106.35	−94.75	−41.36	−22.51	−87.51	−99.54	−118.16
N2—C3—C4—O5	174.50	173.51	−1.41	−9.11	69.35	35.50	−25.28	66.58	−70.32	−57.43
C4—C3—C6	111.30	111.56	111.54	111.26	113.86	112.28	109.44	114.42	113.92	112.65
N2—C3—C6—C7	−13.45	31.83	22.17	−29.45	11.30	17.92	−10.86	−37.19	15.34	−19.90

(Continued)

TABLE I
(Continued)

	1	2	3	4	5	6	7	8	9	10
E	0.000	2.264	3.104	4.448	6.353	8.348	8.577	8.974	9.197	10.682
C3—C6—C7—C8	32.97	−37.25	−36.84	38.23	−31.52	7.52	31.48	38.11	−33.23	34.12
O5—C3—C4—O9	−179.69	−179.11	179.90	−178.67	−179.77	176.93	−178.38	179.50	179.78	−179.16
C6—C3—H10	110.27	110.65	110.92	109.80	111.04	111.76	110.64	110.13	111.27	110.83
O9—C4—O5—H11	−0.25	−0.09	−179.63	−177.06	3.10	−0.36	0.39	2.76	−2.43	−2.42
C6—C7—C8—H12	79.28	151.11	157.88	84.52	159.16	89.83	79.42	93.02	158.16	80.87
C6—C7—C8—H13	−159.79	−88.68	−81.18	−154.76	−79.74	−149.35	−159.35	−146.88	−81.02	−158.00
C3—C6—C7—H14	−83.85	−159.01	−158.31	−78.98	−153.33	−110.32	−85.38	−79.44	−154.71	−82.90
C3—C6—C7—H15	154.84	80.12	80.51	159.96	85.27	129.32	153.32	160.04	83.73	156.13
N2—C3—C6—H16	103.96	155.45	145.71	87.86	134.11	139.79	106.85	80.47	138.63	97.58
N2—C3—C6—H17	−136.41	−84.97	−94.68	−152.53	−106.19	−101.33	−133.81	−159.42	−101.95	−142.40

H—O—C=O, which, as already mentioned, can take values around 0° and 180°; second, one of the torsions N—C—C—O, which define the orientation of the carboxyl group relative to the ring; third, one of the torsions N2—C3—C6—C7 or C3—C6—C7—C8 to describe the pucker-up or pucker-down orientation of the pyrrolidine ring [31] (which are also known as C^{γ}-*exo* and C^{γ}-*endo* [32], or N and S [33], respectively); and, finally, the orientation of the amino group hydrogen atom that may occupy two different positions relative to the substituents of the asymmetric carbon atom C3.

The 6-311 + +G** results given here were obtained in a stepwise fashion: First, the PES was searched with a small split-valence basis set (4-31G). The results obtained in this search were then used as initial data in 6-31G* calculations, and the results of these were used as the initial values for the 6-311 + +G** calculations. This stepwise technique, the primary intent of which was an economic use of computational facilities, revealed several interesting basis-set dependencies. One of these regards the conformer pair **1**/**2**: With the 4-31G basis set, the relative stabilities of **1** and **2** are interchanged, i.e., **2** is the global minimum with this basis set and **1** has a relative energy of 1.497 kJ/mol. With the 6-31G* basis set, **1** is more stable by 1.614 kJ/mol, and with the 6-311 + +G** basis set, **1** is more stable by 2.264 kJ/mol. This change in relative energy is accompanied by changes in the optimized structures; the largest deviations occur for the torsion angle H1—N2—C3—C4, which has the following optimized values in **1**: −26.15° with 4-31G, −21.96° with 6-31G*, and −16.57° with 6-311 + +G**. The vibration spectra analysis of proline and hydroxyproline, which was mentioned above, is based on 4-21G-optimized geometries; in the case of proline, an optimized value of −41.38° is reported [11] for the H1—N2—C3—C4 torsion. Since these values show a remarkable trend, we also optimized the structures of **1** and **2** with several other standard basis sets to check the adequacy of our basis-set choice. The resulting values for the torsion H1—N2—C3—C4 and the energy difference between conformers **1** and **2** are displayed in Figure 3. These values show clearly that in the present case the inclusion of diffuse functions is of minor importance for the energetics, but of significant importance for structural details.

Another type of basis-set dependence occurs for conformer **8**: This conformer is a true local mini-

mum in the 4-31G PES, but a stationary point of inflection with the basis sets 6-31G* and 6-311 + + G**. Such stationary points can be considered as local minima that have one reaction path without a potential barrier [34]; in the case of **8**, it is the reaction leading to **5**.

Still another type of basis-set dependence concerns the conformation with torsion angles H1—N2—C3—C4 = − 42.95°, N2—C3—C4—O5 = 44.37°, N2—C3—C6—C7 = 0.57°, and O9=C4—O5—H11 = 0.12°. In the 4-31G PES, this conformation is a true local minimum (**11**) with distinct saddle points in reaction paths to **1**, **6**, **7**, and **8**; with either 6-31G* or 6-311 + +G**, however, no local minimum exists with a similar conformation. All geometry optimizations starting in that region of the PES converge to **6**. Calculations including electron correlation via Møller–Plesset second-order perturbation theory show the same behavior: Two local minima (**6** and **11**) exist with the 4-31G basis set, but only one (**6**) with the 6-31G* and the 6-311 + +G** basis set.

The reaction paths, which interconnect the 10 local minima described above, are listed in Table II. In all reaction paths, except those that involve an internal rotation of the —OH group, the internal rotation of the —COOH group is coupled either with an inversion of the amino group or with a ring pucker motion. Figures 4 and 5 show characteristic examples of this coupling, which is

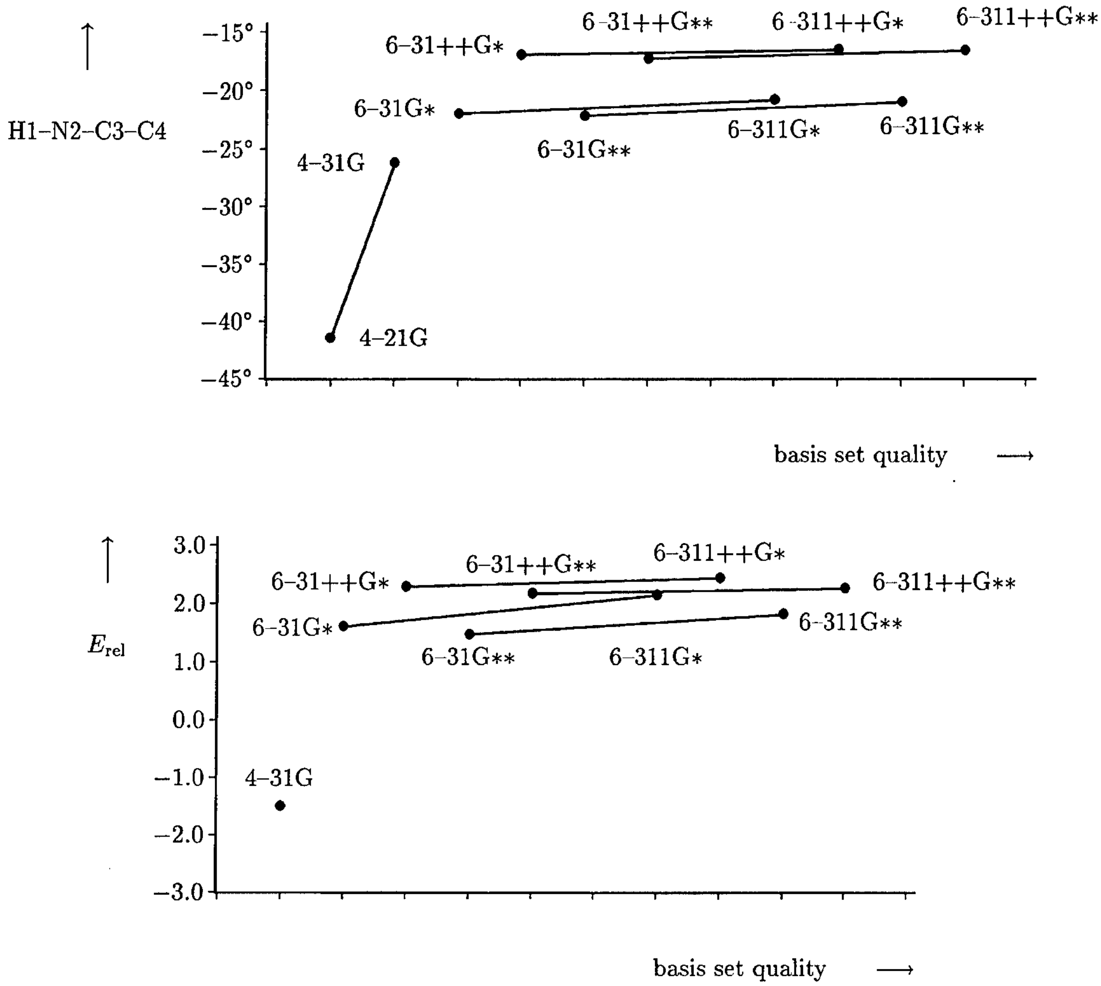


FIGURE 3. Basis-set dependence of the optimized value of the H1—N2—C3—C4 torsion in conformer **1** (top) and the energy difference between conformers **2** and **1** (bottom). Values for basis sets with an identical number of polarization and diffuse functions are connected by straight lines; basis sets are ordered (on a nonlinear scale) according to decreasing RHF energy.

TABLE II
Reaction paths and potential barriers (kJ / mol) in the RHF / 6-311 + +G PES of L-proline.**

Initial conformer	Final conformer	Reaction path description	Barrier
1	**2**	*e*	4.11
1	**6**	*c* and *e*	13.29
2	**1**	*e*	1.84
2	**5**	*c* and *f*	11.42
2	**9**	*d* and *f*	13.76
3	**4**	*e*	8.17
3	**5**	*a* and *d*	44.96
3	**9**	*b* and *c*	45.66
4	**3**	*e*	6.87
4	**10**	*b* and *c*	44.80
4	**8**	*a* and *d*	46.58
5	**2**	*d* and *f*	7.15
5	**3**	*a* and *d*	41.74
5	**6**	*f*	13.52
5	**8**	*e*	3.52
6	**1**	*c* and *e*	4.94
6	**5**	*f*	11.52
6	**7**	*c* and *e*	0.79
6	**8**	*e* and *f*	13.83
6	**9**	*c* and *f*	13.74
7	**10**	*f*	14.48
7	**6**	*d* and *e*	0.56
8	**4**	*b* and *c*	42.05
8	**5**	*e*	—
8	**6**	*e* and *f*	13.20
9	**2**	*c* and *f*	6.82
9	**3**	*a* and *d*	41.94
9	**6**	*d* and *f*	12.89
9	**10**	*e*	6.94
10	**4**	*a* and *d*	38.56
10	**7**	*d* and *f*	12.37
10	**9**	*e*	5.45

Internal rotations are labeled as follows—*a*: decrease of torsion H11—O5—C4—C3; *b*: increase of torsion H11—O5—C4—C3; *c*: decrease of torsion N2—C3—C4—O5; *d*: increase of torsion N2—C3—C4—O5; *e*: ring pucker motion; *f*: amino group inversion.

caused by attractive interactions N—H···O=C (**1** and **2**) and N—H···O—C (**6** and **7**). The H···O distances are 2.391, 2.353, 2.379, and 2.472 Å in **1**, **2**, **6**, and **7**, respectively.

The bond orders [35] of these N—H···O interactions are less than 0.01 in all four cases, nor do the calculated N—H vibration frequencies show a significant deviation pattern; the N—H···O interactions, hence, cannot be classified as true hydrogen bonds. They are, however, strong enough to prevent the existence of four analogous conformers with an inverted amino group and thus explain the difference between the actual number of conformers with *syn*-periplanar orientation of the groups C=O and O—H, which is 8, and an estimate based on the number of degrees of freedom for the individual torsions (two for puckering, two for amino group inversion, three for internal rotation of the —COOH group), which anticipates $2 \times 2 \times 3 = 12$ structures.

MM3 Results

The torsional energy of the MM3 force field, which turned out to be the largest energy component for neutral L-proline, is defined as a sum of terms:

$$E_T = \frac{V_1}{2}(1 + \cos\omega) + \frac{V_2}{2}(1 - \cos 2\omega) + \frac{V_3}{2}(1 + \cos 3\omega),$$

with ω being the dihedral angle for four bonded atoms A—B—C—D and V_1, V_2, and V_3 being the first-, second-, and third-order torsional constant, respectively. Similar expressions are also used in other force fields (e.g., MMFF94 [36] and MSX [37]); hence, the following discussion is of relevance beyond the specific case of MM3(94). The nitrogen atom of the pyrrolidine ring is in a pyramidal conformation, so the torsional angle N—C—C—O is of type 8-1-3-75 in the MM3 nomenclature. This torsional parameter type is not present in amides or peptides [38–40], and, obviously, for that reason, MM3(94) has not yet been parametrized for this type of torsion.

Different sets of parameters were used to locate the minima of the proline PES. The parameter sets used are collected in Table III, and the resulting positions of the local minima are shown in Figures 6 and 7. The first two parameter sets, labeled I and II in the following, were straightforward attempts based on the strategy that the interesting parameter can be compared with other parameters for the same central atoms A—B—C—D, because the nature of terminal atoms A and D is considered to be of minor influence in building a model compound for force field parametrization [41]. The types 9-1-3-75 [i.e., N(sp^2)—C—C—O], 8-1-3-9 [i.e., N—C—C—N(sp^2)], and 8-1-3-7 (i.e., N—C—C=O) have all parameters equal to zero, so $V_1 = V_2 = V_3$

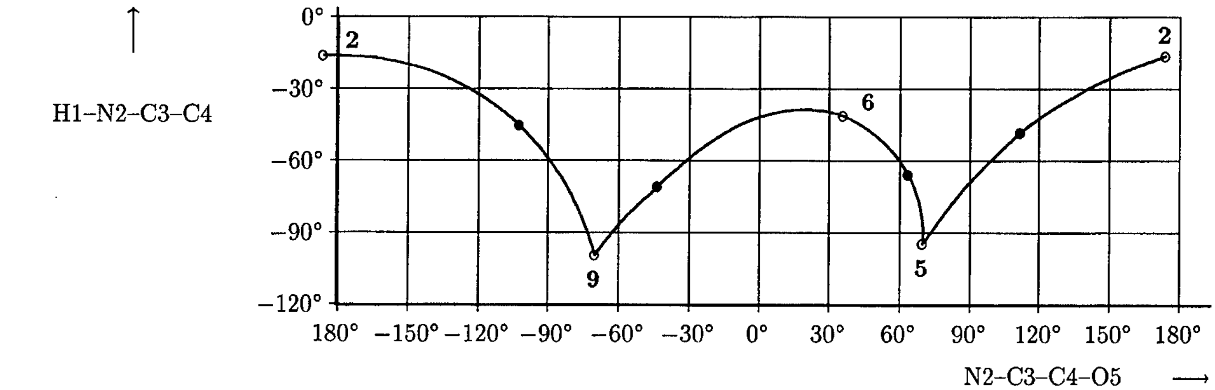


FIGURE 4. The inversion of the amino group is coupled with the internal rotation of the —COOH group through intramolecular N—H ··· O interactions in **1** and **7** as well as **6** and **2**, as shown here for the reactions involving the latter pair. (○) Local minima; (●) transition states.

= 0 was our choice for parameter set I. Set II was chosen from the parameters for nitrogen atoms with similar behavior, namely, sp^3 hybridization, so it employs the data for type 39-1-3-7 (i.e., N^+—C—C=O) [42]. Parameter set III is based on the estimation routine, which is implemented in MM3(94) for undefined parameters [43]. This routine performs ad hoc comparisons with torsional barriers of fragments with the same central atoms B—C and similar geometrical behavior.

Eight minima with *syn*-periplanar groups C=O and O—H could be located with all of these three parameter sets, whereby the parameter choice had almost no effect on the optimized structures in most cases. Comparison of these local minima with the ab initio conformers shows that only two of them are similar in structure. These comparable structures are **1** and **8**. For **8**, the MM3-optimized H—N—C—C dihedral angle is −71.7°, which is 16° larger than the ab initio value; for **1**, the N—C—C—O torsion angles differ by 20° and the H—N—C—C torsions differ by 8°;

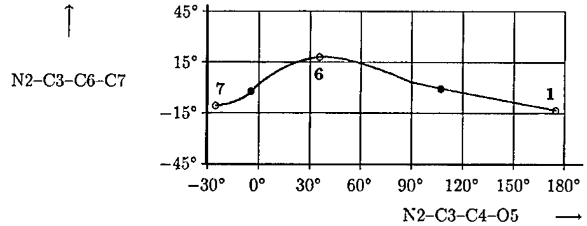


FIGURE 5. The ring-puckering motion is coupled with the internal rotation of the —COOH group in many reaction paths, as shown here for the reactions **7** ⇌ **6** ⇌ **1**. (○) Local minima; (●) transition states.

and the other torsional parameters differ less than 3° in both cases. With all three parameter sets, the global minimum has a geometry with dihedral angles H—N—C—C = −39.7° and N—C—C—O ≈ 0° (70.7° with parameter set I, 69.5° with set II, and 68.9° with set III), the pyrrolidine ring being in a pucker-down orientation. This geometry is somehow similar to the ab initio conformer **6**, but the N—C—C—O torsion is about 35° too large with all three parameter sets; thus, the global minimum is not really comparable to **6**. The other geometries are not comparable with any of the ab initio structures, as documented in Figure 6.

Since these straightforward parameter guesses are all equally unable to describe at least the global minimum correctly, the necessity of a more elaborate estimate for the parameter N—C—C—O on the basis of ab initio results is obvious. In the case of proline, this also poses problems, because the internal rotations of the carboxyl group are all coupled with amino group inversions or ring pucker motions; hence, a "pure" N—C—C—O rotation profile cannot be extracted. Calculating the missing torsional parameter via least square fit of energy profiles of a model compound containing this specific torsional angle is a well-known method [41]. The 6-31G** basis set is regarded to be sufficient to fit the MM3 force-field geometries [44]. Since the smallest neutral amino acid glycine is not suitable because of its symmetrical energy profiles, the energy profiles of the pure N—C—C—O rotation of alanine [45] between conformers 1, 4, 5, 6, 7, 8, and 9 of [7] were used to fit the data of parameter set IV, which are listed in Table III.

TABLE III
Torsional constants (kcal / mol) used in the MM3 calculations for torsion type 8-1-3-75.

Parameter set	V_1	V_2	V_3
I	0.000	0.000	0.000
II	0.000	0.000	0.100
III	0.000	0.160	0.090
IV	0.398	3.138	0.927

With this parameter set IV, we were able to locate 12 local minima with *syn*-periplanar orientation of the groups C=O and O—H, the positions of which are shown in Figure 7. The structure of MM3 global minimum agrees well with that of the RHF global minimum, and there are more structures that are similar to RHF conformers within a 25° tolerance in the torsion angles: Similarities occur for conformers **1**, **2**, **6**, **7**, **8**, and **10**. In the case of **2**, the MM3 structure has, however, an almost planar ring, and in the case of **6**, there are two MM3 structures that differ in the N2—C3—C6—C7 torsion by approximately 15°. The remaining six minima, however, do not match with RHF conformers and also the coupling of internal coordinates in many reaction paths is not reproduced with this force field. Regarding the hydrogen-bonded conformers **3** and **4**, parameter set IV is capable of locating one structure with a N—H distance of 2.18 Å, which deviates from **3** in the N—C—C—O torsion about 50°. Hence, parameter set IV also does not lead to respectable results.

Discussion

One item of interest is certainly the intramolecular N···H—O hydrogen bond in conformers **3** and **4**. According to the RHF/6-311++G** results, the distances of these hydrogen bonds are 2.003 Å in **3** and 2.032 Å in **4**, with bond orders of 0.019 and 0.030, respectively. These data give an unclear picture of the relative strength of the hydrogen bonds: According to the distance, the H bond in **3** is stronger, whereas according to the bond order, the one in **4** is stronger. Elongation of the O—H bond and change in the calculated O—H

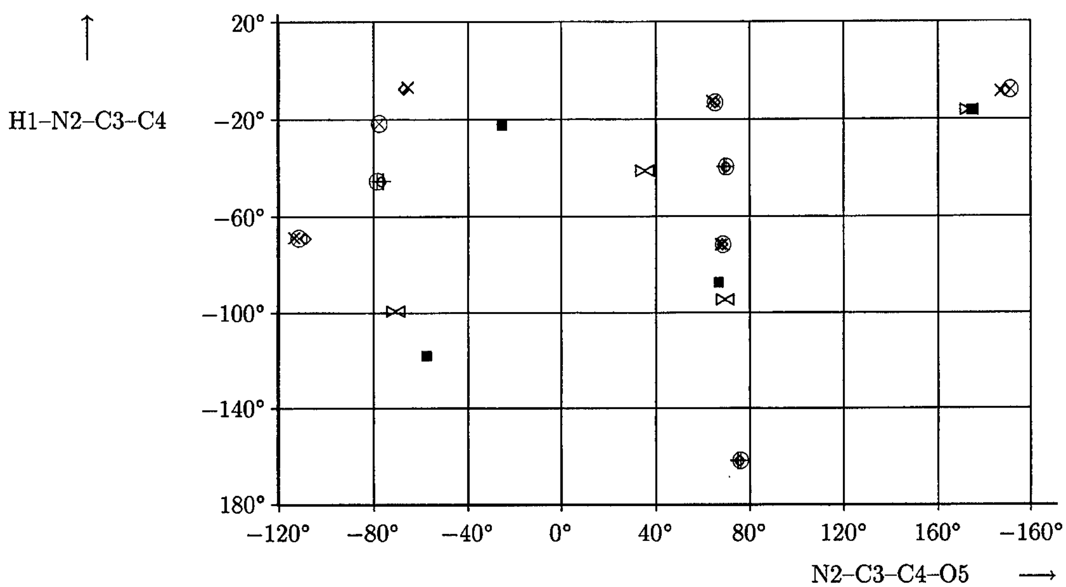


FIGURE 6. Positions of the local minima with *syn*-periplanar groups C=O and O—H according to RHF / 6-31++G** and MM3(94) calculations with parameter sets I, II, and III. The pucker-up positions are labeled (⊗) parameter set I, (◇) set II, (×) set III, and (■) RHF; the pucker-down positions are labeled (⊕) set I, (○) set II, (+) set III, and (⋈) RHF. The positions of the global minima are N2—C3—C4—O5 = 174.4°, H1—N2—C3—C4 = −16.57° (RHF), and N2—C3—C4—O5 ≈ 70°, and H1—N2—C3—C4 = −39.7° (MM3, parameter sets I, II, and III).

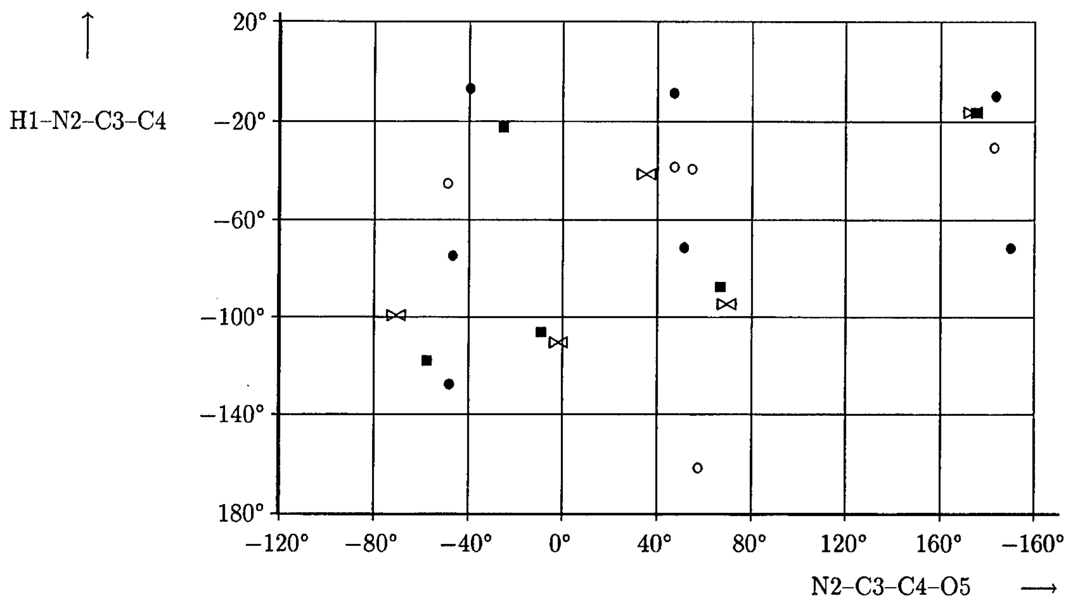


FIGURE 7. Positions of the local minima with *syn*-periplanar groups C═O and O—H according to RHF / 6-311 + + G** and MM3(94) calculations with parameter set IV: (●) MM3; (■) RHF, ring pucker-up; (○) MM3; (⋈) RHF, ring pucker-down. The positions of the global minima are N2—C3—C4—O5 = 174.4°, H1—N2—C3—C4 = −16.57° (RHF), N2—C3—C4—O5 = −176.8°, and H1—N2—C3—C4 = −9.9° (MM3).

frequency (3991.8 cm^{-1} in **3** and 4004.1 cm^{-1} in **4**, which compare to a mean value of 4118.0 cm^{-1} for conformers **1**, **2**, **5**, **6**, **7**, **8**, **9**, and **10**) both indicate the H bond in **3** to be stronger. A different judgment of H-bond strength according to various criteria was observed earlier for ω-amino acids [46, 47] and ω-hydroxy acids [48]. However, in these cases, the H-bonded cycles were eight- or nine-membered and, hence, sterically much more complex than in the present case. The present case of a five-membered, almost planar cycle closed by the H bond is, however, restricted by the pyrrolidine ring of the proline molecule. This restriction is, e.g., manifested by the potential barriers of the ring pucker motion, which are highest in the pair **3** and **4** (cf. Table II) and is the most likely reason for the distance/bond-order discrepancy in proline since a possible electronic interaction C7—H···O in **3** can be ruled out due to the distance of 3.4 Å between these atoms. A comparison with those forms of glycine and α-alanine, which contain the same H-bonded cycle, should give more insight. Geometries of glycine and α-alanine, which were optimized at RHF and MP levels with the same basis set as used in the present study, are described in [6]; unfortunately, the distances between the hydrogen-bonded atoms are not listed in that reference.

A repetition of these optimizations gave the following results for the intramolecular N···H—O hydrogen bond: in glycine, a distance of 2.063 Å (the same value as obtained by Hu et al. with a TZ2P basis set [4]) and a bond order of 0.019; and in alanine, a distance of 2.030 Å and a bond order of 0.023. In these data, a bond length decrease is accompanied by an increase of bond order, so these two measures for H-bond strength agree well. However, the stronger bond occurs for alanine, not or glycine, although glycine definitely has minimal sterical strain. The increase of H-bond strength when going from glycine to α-alanine may be interpreted either as an electronic effect of the methyl group onto the lone pair of the nitrogen atom or as a steric effect of pushing the —COOH group away from the side chain and thus closer to the nitrogen atom. In view of the data for **3** and **4**, it appears that the explanation as a steric effect is correct, because in **4**, the CH_2—C═O fragment is in a staggered orientation, whereas in **3**, it is eclipsed, which adds another push to the —COOH group.

A different topic that deserves discussion is the comparison of the amino acid as an isolated molecule with the amino acid as the building block of peptides and proteins. Such a comparison can nicely be done for proline, because a theoretical

structure investigation of *N*-acetyl-L-proline amide has been published recently [49]. A three-letter code *xYz* was used to label the amide conformers, in which $x = c, t$ denotes the orientation of the acetyl group (*cis* vs. *trans*), $X = A, B, C, F$ describes the orientation of the—$CONH_2$ group relative to the pyrrolidine ring in the notation introduced by Zimmerman et al. [50], and $z = u, d$ indicates the orientations pucker-up and pucker-down. In the notation of [50], A characterizes peptide conformations with $-110° \leq \phi \leq -40°$ and $-90° \leq \psi \leq -10°$, B characterizes conformations with $-180° \leq \phi \leq -110°$ and $-40° \leq \psi \leq 20°$ or $-110° \leq \phi \leq -40°$ and $-10° \leq \psi \leq 50°$, C characterizes conformations with $-110° \leq \phi \leq -40°$ and $50° \leq \psi \leq 130°$, and F characterizes conformations with $-110° \leq \phi \leq -40°$ and $\psi \geq 130°$ or $\psi \leq -140°$.

In comparing the results of the free acid with those of the amide, similarities as well as differences can be noted. The similarities are mostly caused by the five-membered pyrrolidine ring and its ability to adopt a pucker-up and a pucker-down orientation and by the relatively low barrier for the up $\rightleftharpoons$ down interconversion. In both systems, there are conformer pairs *Fd/Fu*, *Cd/Cu*, and *Ad/Au* and a conformer *Au* that has no matching *Ad*. In addition, for the free acid, there is a conformer *Bd* (**6**) without a matching *Bu* in the 6-311++G** PES. The most obvious differences are related to the amino group substituent: In the amide, the acetyl group is present either in *cis* or *trans* orientation and the substitution pattern of the nitrogen atom is planar; in the free acid, the character of the nitrogen atom is pyramidal, as already mentioned.

In the amide, the pucker-down conformer is always more stable than is the pucker-up analog, with (6-31G*) energy differences of 3.7, 4.6, and 8.0 kJ/mol; in the free acid, this is also true except for the global minimum **1**, which is more stable than is the corresponding pucker-down conformer **2**. The energy difference between pucker-up and pucker-down is significantly smaller in the free acid, with a maximum value of 2.6 kJ/mol. Inspection of the molecular structures shows that the higher pucker-up–pucker-down energy differences in the amide are caused by repulsive steric effects due to the acetyl group, which is much more space-filling than is the hydrogen atom in the free acid. The largest energy difference, 8. 0 kJ/mol, occurs between the amide conformers *tCd* and *tCu*, which are lowest in absolute energy because they are stabilized by a γ-turn N—H···O=C hydrogen bond that is formed between the amide and the acetyl group. This γ-turn hydrogen bond is characterized by a bond order of more than 0.04 and results in a considerable stabilization; e.g., the (6-31G*) potential barriers of those reactions, which break this hydrogen bond, are 55.6, 75.9, 78.3, 83.0, 90.2, and 97.0 kJ/mol in the case of *tCd*. As a consequence of this hydrogen bond, the fragment H_2C—CO—N—CH_2 is forced into an unfavorable orientation: In *tCd*, the hydrogen atoms in this fragment are orientated next to each other in an eclipsed position, whereas in *tCu*, one of the ring hydrogen atoms points almost directly toward one of the acetyl group hydrogen atoms. The energy difference of 8.0 kJ/mol between *tCd* and *tCu* therefore is not only, to a minor extent, an effect of the pyrrolidine ring puckering, but, rather, also a consequence of the γ-turn hydrogen bond.

Another interesting aspect of the comparison between free acid and amide is the energetical order of the conformers as a function of the orientation of the —CO—R group. In the amide, this order is $C < B \approx A < F$; in the free acid, it is $F < C \approx B \approx A$. Comparing these orders, one notes that conformation C is more stable in the amide, which is a direct consequence of the stabilizing γ-turn N—H···O=C hydrogen bond discussed above. One also notes that conformation F is shifted to significantly lower energies in the free acid: In the amide, the relative (6-31G*) energies of conformers *cFd* and *cFu* are 24.008 and 27.703 kJ/mol, respectively, whereas in the free acid, the *Fu*-conformation (**1**) is the global minimum and *Fd* (**2**) has a relative energy of 2.264 kJ/mol. This shift in energy (conformers highest in energy vs. conformers of lowest energy) is much too large to be a consequence of the weak N—H···O=C interaction in **1** and **2**. There also is no steric effect of the CH_3 terminal of the acetyl group, as the corresponding *N*-formylproline amide conformers show that they do not have this terminal CH_3 group: The relative (6-31G*) energies are 21.931 kJ/mol for *cFd* and 23.810 kJ/mol for *cFu* in this case [51]. Hence, this shift in energy can only be ascribed to the orientation of the C=O group, which corresponds to the peptide torsion $\omega = 0°$ in *cFd* and *cFu* and also in *cBd* and *cAu*. The latter pair does not show such a dramatic shift in energy, which is explainable by the attractive N···H—N—CO interaction that is present in these two conformers. This interaction is of similar strength as the N—H···O=C interaction in the F conformation of the free acid. In total, these individual effects

give the following picture: The 0° orientation of the central peptide bond torsion ω is considerably destabilized; this destabilization is lowered in conformation A and B by attractive interactions in the amide and amplified in the F orientation by attractive interactions in the free acid. A quantitative treatment would be highly desirable but cannot be performed in a straightforward manner, because the absolute RHF energies of the two systems are not comparable.

Summary

The energetically low-lying parts of the PES of neutral L-proline were investigated by ab initio RHF calculations with different basis sets. Various basis-set dependencies could be noted regarding the number and the nature of the local minima as well as a significant trend in some structural parameters; the inclusion of diffusion functions appears to be essential in this case.

Similar calculations were performed using the MM3(94) force field, which lacks constants for the N—C—C—O torsion in neutral amino acids. Straightforward guesses for these constants turned out to be totally misleading. A more elaborate guess based on RHF/6-31G* data of α-alanine gave somehow more realistic results. This guess, however, was also unable to reproduce certain significant features of the RHF PES, so it seems that the behavior of the pyrrolidine ring in L-proline cannot be sufficiently described by calculating only one missing parameter using the open-chain molecule α-alanine.

The comparison of the results obtained for the free acid with those of the model dipeptide compound *N*-acetylproline amide as well as with those of α-alanine and glycine gives interesting insight into stability patterns of structural fragments, which cannot be gathered from either of the individual species.

References

1. H. L. Sellers and L. Schäfer, J. Am. Chem. Soc. **100**, 7728 (1978).
2. L. Schäfer, H. L. Sellers, F. J. Lovas, and R. D. Suenram, J. Am. Chem. Soc. **102**, 6566 (1980).
3. R. D. Suenram and F. J. Lovas, J. Am. Chem. Soc. **102**, 7180 (1980).
4. C.-H. Hu, M. Shen, and H. F. Schaefer III, J. Am. Chem. Soc. **115**, 2923 (1993).
5. S. Gronert and R. A. J. O'Hair, J. Am. Chem. Soc. **117**, 2071 (1995).
6. A. G. Császár, J. Mol. Struct. **346**, 141 (1995).
7. M. Cao, S. Q. Newton, J. Pranata, and L. Schäfer, J. Mol. Struct. (Theochem) **332**, 251 (1995).
8. A. G. Császár, J. Phys. Chem. **100**, 3541 (1996).
9. A.-M. Sapse, L. Mallah-Levy, S. B. Daniels, and B. W. Erickson, J. Am. Chem. Soc. **109**, 3526 (1987).
10. I. D. Reva, S. G. Stepanian, A. M. Plokhotnichenko, E. D. Radchenko, G. G. Sheina, and Yu. P. Blagoi, J. Mol. Struct. **318**, 1 (1994).
11. P. Tarakeshwar and S. Manogaran, J. Mol. Struct. (Theochem) **365**, 167 (1996).
12. L. Piela, G. Némethy, and H. A. Scheraga, Biopolymers **26**, 1587 (1987).
13. R. Sankararamakrishnan and S. Vishveshwara, Biopolymers **30**, 287 (1990).
14. P. Y. Chou and G. D. Fasman, Biochemistry **13**, 211 (1974).
15. P. Y. Chou and G. D. Fasman, Biochemistry **13**, 222 (1974).
16. S. Dasgupta and J. A. Bell, Int. J. Peptide Protein Res. **41**, 499 (1993).
17. M. Ramek, J. Mol. Struct. (Theochem) **208**, 301 (1990).
18. M. Ramek and V. K. W. Cheng, Int. J. Quantum Chem.: Quantum Biol. Symp. **19**, 15 (1992).
19. M. Ramek and M. Flock, Amino Acids **8**, 271 (1995).
20. M. Flock and M. Ramek, Int. J. Quantum Chem.: Quantum Chem. Symp. **26**, 505 (1992).
21. M. Flock and M. Ramek, J. Mol. Struct. (Theochem) **310**, 269 (1994).
22. V. K. W. Cheng, M. Flock, and M. Ramek, Int. J. Quantum Chem. **57**, 929 (1996).
23. U. Seebacher and M. Ramek, Amino Acids **7**, 223 (1994).
24. C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).
25. W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).
26. P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).
27. T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. von Ragué Schleyer, J. Comput. Chem. **4**, 294 (1983).
28. N. L. Allinger, Y. H. Yuh, and J.-H. Lii, J. Am. Chem. Soc. **111**, 8551 (1989). The MM3 program is available to commercial users only from Tripos Associates, 1699 South Hanley Road, St. Louis, MO 63144 and to academic users only from QCPE, Indiana University, Bloomington, IN 47405.
29. M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, Jr., J. Comp. Chem. **14**, 1347 (1993).
30. J.-H. Lii, private communication.
31. F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, J. Phys. Chem. **79**, 2361 (1975).
32. R. Balasubramanian, A. V. Lakshminarayanan, M. N. Sabesan, G. Tegoni, K. Venkatesan, and G. N. Ramachandran, Int. J. Protein Res. **3**, 25 (1971).
33. C. A. G. Haasnoot, F. A. A. M. de Leeuw, H. P. M. de Leeuw, and C. Altona, Biopolymers **20**, 1211 (1981).

34. P. G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987), p. 71 ff.
35. I. Mayer, Chem. Phys. Lett. **97**, 270 (1983).
36. T. A. Halgren, J. Comp. Chem. **17**, 490 (1996).
37. E. Demiralp, S. Dasgupta, and W. A. Goddard III, J. Phys. Chem. **101**, 1975 (1997).
38. D. Q. McDonald and W. C. Still, Tetrahedron Lett. **33**, 7743 (1992).
39. J. Lii and N. L. Allinger, J. Comp. Chem. **12**, 186 (1991).
40. D. Q. McDonald and W. C. Still, J. Org. Chem. **61**, 1385 (1996).
41. A. J. Hopfinger and R. A. Pearlstein, J. Comp. Chem. **5**, 486 (1984).
42. N. L. Allinger, private communication.
43. N. L. Allinger, X. Zhou, and J. Bergsma, J. Mol. Struct. (Theochem) **331**, 69 (1994).
44. J. Lii and N. L. Allinger, J. Phys. Org. Chem. **7**, 591 (1994).
45. A.-M. Kelterer, to be published.
46. M. Ramek, Struct. Chem. **6**, 15 (1995).
47. M. Ramek, Int. J. Quantum Chem.: Quantum Biol. Symp. **21**, 79 (1994).
48. M. Flock, Int. J. Quantum Chem.: Quantum Chem. Symp. **29**, 585 (1995).
49. M. Ramek, A.-M. Kelterer, B. J. Teppen, and L. Schäfer, J. Mol. Struct. (Theochem) **352 / 353**, 59 (1995).
50. S. S. Zimmerman, M. S. Pottle, G. Némethy, and H. A. Scheraga, Macromolecules **10**, 1 (1977).
51. Unpublished data; the conformers *tCu* and *tAu* of N-formylproline amide are discussed in a different context in A.-M. Kelterer, M. Ramek, R. F. Frey, M. Cao, and L. Schäfer, J. Mol. Struct. (Theochem) **310**, 45 (1994).

Correlation of Ultraviolet Spectra with Structure via the Integrated Molecular and Electronic Transforms

STEPHEN P. MOLNAR,[1] JAMES W. KING[2]

[1] *Foundation for Chemistry; 1945 Elmwood Avenue, Upper Arlington, Ohio 43212*

[2] *Foundation for Chemistry, P.O. Box 116, Balsam, North Carolina 28707*

Received 1 March 1997; accepted 23 June 1997

ABSTRACT: The integrated molecular transform (FT_m) has been used for the correlation of the structures of organic molecules with their physicochemical, thermodynamic, and pharmacological properties; it is also an excellent conformation index and functions as a discriminator of classical chemical structure types. In this study, it is used along with our recently introduced normalized molecular moment (M_n), and new structure indices, viz the integrated electronic transform (FT_e), the integrated charge transform (FT_c), and the electronic moment (M_e), to establish appropriate models for the title subject. Initially, the principal absorption maxima in each of several series were regressed against the structural indices to determine which index best represented the structures in the context of the absorption data. The indices were then selectively regressed against the absorption data to generate absorbance estimation equations. In a series of multicyclic hydrocarbons, the FT_m functioned as a topological structure discriminator as well as a structure surrogate. In the topological subsets, the FT_e and FT_c also were selectively useful. For a series of conjugated dienes, the FT_m and the M_n were statistically appropriate. In a series of substituted benzenes, the discrimination of halobenzenes was apparent and could be represented by either the FT_m, FT_e, or M_e indices. For other variously substituted benzenes, the FT_m is the extant model and further work with larger, structurally delineated series is warranted. For a series of monoalkyl-substituted nitrobenzenes, the FT_m and FT_e parameters are appropriate variables. Satisfactory correlation of molar absorptivities was not possible in this study as it would require absorption curve integration in the range where the maxima occurs. © 1997 John Wiley & Sons, Inc. Int J Quant Chem **65**: 1047–1056, 1997

Correspondence to: J. W. King.

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061047-10

Introduction

The integrated molecular transform (FT_m) is derived from the molecular transform [$I(s)$] of Soltzberg and Wilkins [1, 2], which in turn was based on the electron diffraction analysis work of Wierl [3]. In this methodology, molecular descriptors such as bond lengths, bond counts, or other parameters were treated by a Fourier transform weighted by the atomic numbers of the constituent atoms of a molecule, to give ordinates of a curve; the curve was then used to generate a 100-digit binary number which could be used as a structure surrogate in structure-activity/property correlation and prediction studies. In our formulation, which utilizes interatomic bond distances [Eq. (1)], the ordinates of the $I(s)$ curve are squared, the area under the resulting curve integrated, and the square root of the area taken as the FT_m, as shown in Eq. (2).

$$I(s) = \sum_{i=2}^{n} \sum_{j=1}^{i-1} \frac{A_i A_j}{s r_{ij}} \sin \left[\left[\begin{matrix} x_i \\ y_i \\ z_i \end{matrix} \right]^2 - \left[\begin{matrix} x_j \\ y_j \\ z_j \end{matrix} \right]^2 \right]^{0.5}, \quad (1)$$

$$FT_m = \sqrt{\int_1^{31} I^2(s)\, ds}\,. \quad (2)$$

A_i and A_j in Eq. (1) are form (or weighting) factors. In the FT_m case, they are the atomic numbers of the atoms between which the coordinate distance was determined in either a two-center or multicenter aspect. In this work, we introduce additional transforms, i.e., the integrated molecular electronic transform (FT_e) in which the weighting factors are the calculated electron densities of the respective atoms and the integrated charge transform (FT_c) in which the weighting factors are the calculated atomic charges on the respective atoms.

We have recently introduced [4, 5] another structure index, the normalized molecular moment (M_n), which is the sum of the products of the atomic weight of each constituent atom multiplied by its distance from the geometric center of the molecule, divided by the molecular weight [Eq. (3)]. In the present instance, we also introduce the molecular electronic moment (M_e), which is the sum of the products of the electron density on each atom multiplied by its distance from the geometric center of the molecule [Eq. (4)].

$$M_n = \sum_{j=1}^{n} \frac{c_j a_j}{w_m}, \quad (3)$$

$$M_e = \sum_{j=1}^{n} \frac{c_j \rho_j}{\sum_{j=1}^{n} \rho} \quad (4)$$

where M_n is the normalized molecular moment, n is the number of atoms in the molecule, c_j is the distance from the molecular geometric center to atom j (Å), a_j is the atomic weight of atom j, W_m is the molecular weight of the molecule, M_e is the normalized electronic moment, and, ρ_j is the calculated electron density of atom j.

In each of these formulations, interatomic distances may be based on standardized bond distances or determined from structures optimized by molecular mechanics methodology or quantum chemical methods (e.g., MOPAC); the latter also serves to generate the atomic electron densities and charges.

The unitary FT_m index has been used to correlate both physicochemical and pharmacological properties, viz the polarizability and local anesthetic activity in a structurally diverse series [6]; an enthalpy function and heats of formation in a series of hydrocarbons, most of which were branched [7]; the enzyme inhibition activity of several series of organophospborus compounds [8]; the octanol/water partition coefficient in other organophosphorus series where the index was shown also to be a structure discriminator [9]; the gas chromatographic (GC) retention indices of a series of mustard [*bis*(2-chloroethyl) sulfide] analogs [10]; pK_a with structure as well as functioning as a structure discriminator [11]; dermal transport rate with classically delineated structures [12]; as a basis for a simple similarity index [5, 13]; a unitary numerical descriptor of structure conformation [14]; and for correlation of diamagnetic susceptibility of organic compounds in a structurally diverse series [15]. The FT_m correlates well with some theoretical linear solvation energy relationships (TLSER) [16]. In the noted examples, linear modeling was resulted in correlation coefficients of greater than 0.9.

In considering additional molecular physicochemical attributes that might be amenable to modeling with molecular transforms, ultraviolet (UV) spectra came to mind. This was because consideration of structural detail often permits rea-

sonably accurate prediction of UV absorption bands and, to an extent, molecular absorptivity. The methodology developed for this is generally referred to as Woodward's or the Woodward-Fieser rules, and Scott's rules for special cases [17–19]. This suggested that our transform and moment concepts might be applied to the direct correlation of molecular structure and UV spectra. This study describes the results of that application.

Methodology

The structures investigated in this study, their associated absorption data, and calculated indices are shown in Tables I, III, V, and VII. The interatomic distance and electronic information necessary for calculation of the transform and electronic moment indices, and distance to constituent atoms from the geometric center of the molecules, as required for calculation of both moment indices, were obtained by first drawing the structures in ChemDraw Pro (v. 3.5). They were then imported into Chem3D Pro (v. 3.5), which permitted optimization by MM2 or MOPAC93 methods for generation of the necessary distances, and electronic information by the latter [20]. The modeling and statistical data were calculated by SYSTAT [21] or with a TI-59 programmable calculator using the routines of Clark [22].

Results and Discussion

AROMATIC HYDROCARBONS

Table I shows the data for 19 compounds. A plot of the major UV absorption band for these versus their calculated FT_m indices revealed, by observation, three distinct groups, as shown in

TABLE I
Absorption maxima [23, 24], integrated molecular transform, and normalized molecular moment indices for aromatic compounds.

No.	Compounds	$\bar{\nu}_{max}$ (kK)^a	ϵ_{max}^b (L / mol cm) ($\times 10^{-3}$)	FT_m^c	FT_m^d	FT_e^e	FT_c^f	M_n^g	M_n^h
1	Benzene	54.50	46.00	130.8006	132.2525	64.36861	0.859055	1.482279	1.480615
2	Naphthalene	45.40	133.00	197.1029	201.1666	95.36857	1.14874	1.986480	1.982338
3	Acenaphthene	43.80	93.00	221.5079	226.4143	106.66258	1.406341	2.108119	2.122805
4	Anthracene	39.00	180.00	284.7887	288.351	135.77252	1.340451	2.558719	2.553458
5	Naphthracene	36.70	190.00	368.5519	374.9043	175.51878	1.487942	3.112254	3.104805
6	Pentacene	32.30	300.00	454.0849	460.3999	214.92817	1.599764	3.703059	3.624926
7	Phenanthrene	39.40	65.50	273.3747	275.7011	129.76932	1.356991	2.46074	2.459848
8	Chrysene	37.20	150.00	354.3647	370.9624	173.72862	1.514039	2.939143	2.970897
9	1,2-Benzanthracene	34.80	113.00	364.4124	370.7791	173.63894	1.509238	3.049184	2.970879
10	3,4-Benzanthracene	35.60	85.00	349.1559	353.8025	165.97919	1.465682	2.855678	2.862941
11	Triphenylene	38.90	150.00	359.8419	365.5476	172.02089	1.563897	2.798721	2.794283
12	Azulene	37.10	47.00	176.9720	180.5548	85.96730	1.162454	2.021417	2.019807
13	1,2-Benzazulene	33.30	60.00	252.1937	256.3687	120.82545	1.406654	2.515555	2.518336
14	Biphenyl	50.50	52.00	234.9134	239.1399	113.92124	1.108196	2.501849	2.509916
15	Fluorene	48.00	43.00	251.7534	255.3227	120.32683	1.333206	2.380029	2.38081
16	2-Phenylnaphthalene	20.00	65.00	322.1034	327.2158	154.47098	1.335809	3.013256	2.967535
17	1,2-Benzfluorene	38.00	70.00	322.1034	333.6507	156.17543	1.519029	3.013256	2.854355
18	2-Phenylazulene	33.10	85.00	293.1007	300.8000	141.98315	1.542411	3.043103	3.111077
19	Indenoazulene	30.90	64.00	308.1031	313.5010	146.90808	1.698667	2.915621	2.937072

[a] Absorption wave number in kiloKaisers (kK).
[b] Molar absorptivity.
[c] Structure optimization by MM2.
[d] Structure optimization by MO.
[e] Integrated molecular electronic transform.
[f] Integrated molecular charge transform.
[g, h] Normalized molecular moment with structure optimization by MM2 and MO, respectively.

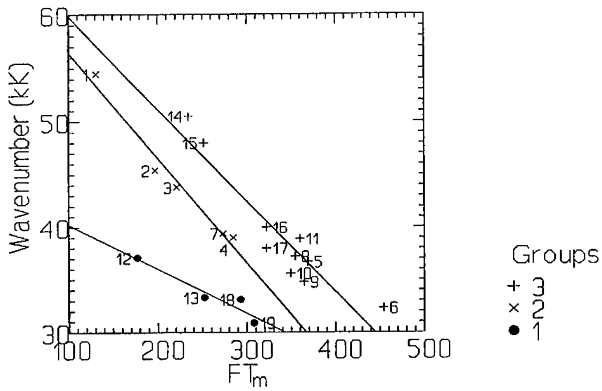


FIGURE 1. FT_m versus absorption wave number (in kiloKaisers) for the aromatic compound series.

Figure 1. Consideration of the structures shows topological similarities within each group that are not necessarily common to the other groups. This example again demonstrates the remarkable power of the FT_m to discriminate classical structure.

Correlation Analysis

Table II shows the results of correlating the principal absorption band and molar absorptivity against the various calculated structural indices for each group. In group 1, the best model for absorption wave number ($\bar{\nu}$) is with the FT_c followed by the FT_m, as judged by the correlation coefficients and F statistics. It is not obvious why the FT_c should stand out in this manner, particularly in comparison to the models for groups 2 and 3. However, this may again reflect the discriminatory power of the transform paradigm, i.e., the group 1 compounds are seven-membered unsaturated rings fused to other ring systems, and this ring system has a greater tendency to form charged carbonium ion species than analogous six-membered rings.

In group 2, the FT_c, while not a bad model for $\bar{\nu}$, is inferior to all the other indices, particularly those involving structure optimization by quan-

TABLE II
Linear correlation of absorption data and molecular indices in Table I.

	$\bar{\nu}_{max}$ and ϵ_{max} vs.					
	FT_m (MM2)	FT_m (MO)	FT_e (MO)	FT_c (MO)	M_n (MM2)	M_n (MO)
	Group 1 (compounds 12, 13, 18, 19)					
$-R_{\nu\ max}$[a]	0.959	0.953	0.948	0.975	0.865	0.845
$F_{\nu\ max}$[b]	22.728	19.619	17.893	38.506	5.924	4.976
$s_{\nu\ max}$[c]	20.450	22.372	10.764	0.062	0.283	0.319
$-R_{\nu\ max}$	0.770	0.782	0.790	0.651	0.895	0.910
$F_{\epsilon\ max}$	2.907	3.149	3.325	1.474	8.013	9.570
$s_{\epsilon\ max}$	45.905	45.839	20.819	0.211	0.252	0.248
	Group 2 (compounds 1, 2, 3, 4, 7)					
$-R_{\nu\ max}$	0.986	0.988	0.987	0.912	0.990	0.991
$F_{\nu\ max}$	104.853	126.038	111.080	14.924	148.102	167.138
$s_{\nu\ max}$	11.995	11.050	5.376	0.107	0.070	0.066
$-R_{\epsilon\ max}$	0.671	0.680	0.678	0.596	0.701	0.638
$F_{\epsilon\ max}$	2.454	2.584	2.562	1.649	2.891	2.057
$s_{\epsilon\ max}$	53.342	53.117	24.347	0.209	0.353	0.380
	Group 3 (compounds 5, 6, 8, 9, 10, 11, 14, 15, 16, 17)					
$-R_{\nu\ max}$	0.933	0.938	0.937	0.882	0.845	0.832
$F_{\nu\ max}$	53.642	59.090	57.106	27.883	19.902	17.951
$s_{\nu\ max}$	23.674	23.082	10.742	0.073	0.205	0.208
$-R_{\epsilon\ max}$	0.882	0.878	0.879	0.664	0.829	0.882
$F_{\epsilon\ max}$	27.916	26.859	27.070	6.294	17.519	28.004
$s_{\epsilon\ max}$	31.014	32.021	14.636	0.116	0.214	0.177

[a] Correlation coefficient.
[b] F statistic.
[c] Standard deviation.

tum chemical (MO) methods as compared to molecular mechanics (MM2).

In group 3, the best models for $\bar{\nu}$ are the FT_m and FT_e with the remainder being very obviously inferior. Again, the discriminatory power of the transforms is evident inasmuch as the compounds in this group are the most extensively conjugated of the three; it would appear that this characterization is reflected in the indices.

In respect to molar absorptivity (ϵ_{max}), none of the indices provide a good model with the lone exception of the normalized molecular moment (M_n), calculated from quantum chemical parameters, for the group 1 compounds. In reality, ϵ_{max} was not well correlated for any series in this study, with the group 1 results being rather typical. This may be due to the fact that, while $\bar{\nu}_{max}$ is usually quite obvious from spectra, ϵ_{max} involves a concentration term and its modeling should reflect the area under the absorption curve because the latter may include a convolution of several less intense absorption, and perhaps, broader, bands. In light of these observations and since actual spectra were not available, ϵ_{max} correlations were not further considered.

CONJUGATED DIENES

Table III shows the absorption maxima and calculated molecular indices for 11 dienes. A plot (not shown) of FT_m against $\bar{\nu}_{max}$ shows a slight curvature due principally to compound 9. This molecule is a bicyclohexenyl with the double bonds unconjugated and is thus slightly different from the remaining compounds in this series.

Correlation Analysis

Table IV shows the data for this series with the M_n being only marginally superior, according to the correlation coefficients and F-statistics, to FT_m and FT_e.

TABLE III
Absorption maxima [25], transform, and moment indices for conjugated dienes.

No.	Compound	$\bar{\nu}_{max}$ (kK)	FT_m	FT_e	FT_c	M_n	M_e
1	1,3-Butadiene	45.870	74.30927	38.62814	0.56652	1.38732	1.51445
2	2,3-Dimethyl-1,3-butadiene	44.043	107.88840	54.72790	0.78878	1.63076	1.78608
3	2,4-Hexadiene	43.849	114.75564	60.40481	0.76457	2.03457	2.19651
4	2-Methyl-1,3-butadiene	45.244	91.02438	46.71057	0.68095	1.49105	1.62970
5	2,4-Pentadiene	44.536	93.26402	48.82182	0.68052	1.70870	1.85779
6	3-(Cyclohexylidene) propene	42.088	165.70787	84.10878	0.89867	2.25838	2.38046
7	2-(1-Cyclohexenyl) propene	42.356	158.92196	79.14550	1.02114	2.13936	2.28263
8	4-(Methylene) isopropylcyclohex-2-ene	42.904	175.84540	88.02970	0.92126	2.30553	2.42180
9	1,3′-Bicyclohexenyl	42.177	243.66184	122.19467	1.21796	2.51875	2.66278
10	1-Cyclohexylidene-2-(4-hydroxycyclohexylidene) ethane	40.136	261.29029	131.93339	1.05001	3.31174	3.42462
11	1-(Methylene) cyclohex-2-ene	43.090	138.76357	69.41557	0.86744	1.72769	1.85094

TABLE IV
Linear correlation of absorption data and molecular indices from Table III.

	$\bar{\nu}_{max}$ vs.				
	FT_m	FT_e	FT_c	M_n	M_e
$-R^a$	0.922	0.922	0.859	0.931	0.601
F^b	50.848	50.684	25.265	58.451	5.092
S^c	25.087	12.416	0.103	0.214	0.560

[a] Correlation coefficient.
[b] F statistic.
[c] Standard deviation.

SUBSTITUTED BENZENES

Table V lists the data for this series. A plot (not shown) of FT_m versus $\bar{\nu}_{max}$ was generally linear but with some clustering evident and with compounds 8 and 20 being apparent outliers, although, as shown below, they could be included in selected correlations. The halobenzenes have classically been regarded as unique and this instance was not an exception.

Correlation Analysis

Table VI shows the data for this series. Once the decision was made to group the halobenzenes by themselves, the correlations fell into place. The model for all compounds except 8 and 20, with correlation coefficients of -0.735 and -0.777 for the FT_m and M_n, respectively, and good F statistics, is probably not sufficiently accurate for predictive use. The halobenzene model is excellent with M_e and FT_m being the better and FT_e next. This order may reflect to some extent the electronic nature of the halogens in this particular series as indicative of their influence on molecular behavior. When other substituted benzenes are added to the model, as in the third part of the table, the correlation coefficients and F statistics drop precipitously except for FT_m with an R value of -0.897 and F statistic of 66.202. These statistics are an indication of the robustness of the FT_m as a unitary index of structure. It would be reasonable to consider use of this model as a preliminary initial predictor of UV absorption bands for series including similar molecules.

NITROBENZENES

Table VII shows the absorption data for this series in both the vapor phase and heptane solution. Again, the highly electronic nature of the nitro substituent, analogous to the halogens, suggested considering these molecules separately. A

TABLE V
Absorption maxima [26] and molecular indices for substituted benzenes.

No.	Compound	$\bar{\nu}$ (kK)	ϵ_{max} ($\times 10^{-3}$) (L/mol cm)	FT_m	FT_e	FT_c	M_n	M_e
1	Benzene	54.4	45.0	132.34663	64.41509	0.85874	1.48022	1.58640
2	Hexaethylbenzene	47.2	40.0	238.54664	119.82010	1.706953	2.83153	3.01723
3	Chlorobenzene	52.7	54.0	136.63966	62.87805	0.75424	1.92053	1.83049
4	*o*-Dichlorobenzene	51.2	60.0	148.49378	61.95596	0.69743	2.15249	1.99793
5	*m*-Dichlorobenzene	51.0	38.0	158.62699	65.43174	0.66886	2.18051	2.02098
6	*p*-Dichlorobenzene	51.8	43.0	166.64601	66.65557	0.64518	2.24461	2.05141
7	1,3,5-Trichlorobenzene	49.4	50.0	204.58498	71.95591	0.56605	2.41777	2.20634
8	Hexachlorobenzene	46.0	90.0	422.80023	100.28110	0.53118	2.66803	2.47292
9	Bromobenzene	52.4	36.0	140.58889	60.02977	0.88729	2.24999	1.85583
10	Iodobenzene	43.9	40.0	155.67567	58.57964	1.07213	2.47846	1.87564
11	Phenol	52.7	50.0	149.42181	72.45640	0.79005	1.67302	1.77013
12	*o*-Cresol	52.2	70.0	157.98198	76.31838	0.89596	1.80292	1.92473
13	*m*-Cresol	51.8	60.0	160.16019	77.82347	0.89545	1.86329	1.98119
14	*p*-Cresol	52.3	45.0	162.61700	79.22689	0.89353	1.88121	2.00826
15	Hydroquinone	52.6	28.0	169.14650	82.26611	0.73114	1.86901	1.96300
16	Aniline	50.8	32.0	147.04542	71.58923	0.93166	1.68716	1.79958
17	Benzoic Acid	51.0	38.0	177.86696	85.33450	0.60347	2.01712	2.10940
18	Benzaldehyde	49.9	26.0	154.42676	74.05752	0.64463	1.89962	1.98488
19	Nitrobenzene	48.5	13.0	183.38723	89.05759	0.23473	2.02554	2.09886
20	Benzonitrile	52.2	44.0	65.47588	31.07088	1.17193	3.35662	3.47191
21	Phenylacetylene	50.2	30.0	150.64614	71.60076	0.87371	1.92476	2.01899
22	Styrene	49.2	23.0	155.34296	75.23977	0.95890	1.86448	1.98530
23	*o*-Toluic acid	50.5	35.0	183.65238	88.08477	0.72782	2.07686	2.18786
24	*m*-Toluic acid	50.0	40.0	192.22253	92.84231	0.66445	2.22185	2.3376

TABLE VI
Linear correlation of absorption data and molecular indices from Table V.

	$\bar{\nu}_{max}$ vs.				
	FT_m	FT_e	FT_c	M_n	M_e
	All compounds except 8 and 20				
$-R$[a]	0.735	0.631	0.142	0.592	0.777
F[b]	23.506	13.265	0.409	10.781	30.422
S[c]	16.956	10.956	0.271	0.249	0.177
	Halobenzenes (compounds 3 to 10 inclusive)				
$-R$	0.956	0.942	0.586	0.801	0.959
F	63.888	47.239	3.146	10.716	68.368
s	30.286	4.902	0.155	0.148	0.066
	Compounds 3 to 17 inclusive and 20, 23, 24				
$-R$	0.897	0.534	0.541	0.343	0.260
F	66.202	6.399	6.610	2.139	1.164
s	29.054	13.581	0.150	0.393	0.387

[a] Correlation coefficient.
[b] F statistic
[c] Standard deviation.

plot (not shown) has a little curvature which is due to nitrobenzene itself. In a way this is not too surprising since the remaining five compounds are substituted with aliphatic groups. In fact, the magnitude of the FT_m index increases with side chain size, an observed characteristic of this index which is paralleled by the FT_e data but not the other indices. Further, superimposing the nitrobenzene plot on that for the substituted benzenes shows that although they are parallel, they are far enough apart that they cannot be included in the same regression. Again, the structural discrimination power of the FT_m is evident.

Correlation Analysis

Table VIII shows the statistical parameters for the substituted nitrobenzenes. Again, the FT_m index is shown to be the best model for absorbancies in the vapor phase as well as in solution. It is followed closely by the FT_e and the M_e. This indicates that, again, the FT_m is the most applicable general index but also that electronic characterization of the molecules can be extracted by the two electronic indices. It is interesting to note that the charge index (FT_c) correlation rules out charge as a factor in describing these molecules, at least in

TABLE VII
Absorption maxima [25] and molecular indices for nitrobenzenes.

No.	Compound	Vapor Phase $\bar{\nu}$ (kK)	Heptane Solution $\bar{\nu}$ (kK)	FT_m	FT_e	FT_c	M_n	M_e
1	Nitrobenzene	41.630	39.452	183.38096	89.05209	0.23403	2.02634	2.09958
2	*p*-Methylnitrobenzene	39.783	37.689	200.88976	98.24634	0.28060	2.23553	2.33137
3	*p*-Ethylnitrobenzene	39.656	37.561	218.36509	108.05060	0.26936	2.45262	2.55669
4	*p*-Propylnitrobenzene	39.562	37.491	238.61608	119.09394	0.13905	2.79221	2.88679
5	*p*-Isobutylnitrobenzene	39.421	37.378	249.31274	124.45922	0.15166	2.89931	2.98748
6	*p*-Tertbutylnitrobenzene	39.312	37.224	252.00474	125.02069	0.32367	2.64511	2.75585

TABLE VIII
Linear correlation of absorption data and molecular indices from Table VII.

	$\bar{\nu}_{max}$ vs.				
	FT_m	FT_e	FT_c	M_n	M_e
		Vapor phase			
$-R$[a]	0.829	0.829	0.015	0.793	0.811
F[b]	8.789	8.814	0.001	6.776	7.678
s[c]	17.359	9.247	0.083	0.228	0.223
		Heptane solution			
$-R$	0.825	0.825	0.003	0.784	0.802
F	8.540	8.540	0.000	6.374	7.228
s	17.531	9.347	0.083	0.233	0.227

[a] Correlation coefficient.
[b] F statistic.
[c] Standard deviation.

the ground state. As noted above, the correlations are influenced by nitrobenzene itself. When this compound is omitted, the correlation coefficients for the substituted compounds rise to -0.962 and -0.935 for the vapor and solution phase absorbancies, respectively, with analogous F statistics of 36.898 and 20.700.

UV ABSORPTION WAVE NUMBER ESTIMATION

In the models discussed above the wave number of UV absorption was regressed, as the independent variable, against the respective calculated structure indices as the dependent variable. This was done in order to elucidate the index most favorable for use in estimation equations. For the latter purpose the variable dependency must be reversed, as one would wish to predict an absorption frequency based on knowledge of structure and a corresponding calculated index. In the following equations this has been done for the reported examples and classes of compounds. In each instance: $\bar{\nu}$ is the absorption wave number, the molecular indexes are as defined above, n is the number of compounds in each regression, R is the correlation coefficient, s is the standard deviation, and F is the F statistic.

Aromatics

Estimations in this series for compounds not included in the original model would require a judgment as to which reported topological group the unknown compound should belong. A preferable way would be to compare the respective FT_m values for similarity, as has been reported [5, 13] and then make the assignment; this would be similarly applicable to the other series in this study.

Group 1

$$\bar{\nu} = -0.042\,FT_m + 44.428,$$
$$n = 4, \qquad R = -0.959,$$
$$s = 0.897, \qquad F = 22.728.$$

Group 2

$$\bar{\nu} = -0.099\,FT_m + 66.426,$$
$$n = 5, \qquad R = -0.986,$$
$$s = 1.209, \qquad F = 104.853.$$

Group 3

$$\bar{\nu} = -0.107\,FT_m + 74.827,$$
$$n = 9, \qquad R = -0.960,$$
$$s = 1.662, \qquad F = 81.556.$$

Conjugated Dienes

$$\bar{\nu} = -0.024\,FT_m + 46.919,$$
$$n = 11, \qquad R = -0.922,$$
$$s = 0.667, \qquad F = 50.848,$$

or

$$\bar{\nu} = -2.737M_n + 28.902,$$
$$n = 11, \qquad R = -0.931,$$
$$s = 0.628, \qquad F = 58.451.$$

Substituted Benzenes

Halobenzenes

$$\bar{\nu} = -0.022\,\mathrm{FT}_m + 54.878,$$
$$n = 8, \quad R = -0.956,$$
$$s = 0.684, \quad F = 63.888,$$

or

$$\bar{\nu} = -0.151\,\mathrm{FT}_e + 61.049,$$
$$n = 8, \quad R = -0.942,$$
$$s = 0.734, \quad F = 47.239,$$

or

$$\bar{\nu} = -9.660 M_e + 70.433,$$
$$n = 8, \quad R = -0.959,$$
$$s = 0.663, \quad F = 68.368.$$

Other substituted benzenes (Table V, compounds 3–17 inclusive and 20, 23, and 24)

$$\bar{\nu} = -0.023\,\mathrm{FT}_m + 55.235,$$
$$n = 18, \quad R = -0.897,$$
$$s = 0.734, \quad F = 66.202.$$

Alkyl-Substituted Nitrobenzenes

Vapor Phase

$$\bar{\nu} = -0.008\,\mathrm{FT}_m + 41.454,$$
$$n = 5, \quad R = -0.962,$$
$$s = 0.059, \quad F = 36.893,$$

or

$$\bar{\nu} = -0.015\,\mathrm{FT}_e + 41.306,$$
$$n = 5, \quad R = -0.950,$$
$$s = 0.067, \quad F = 27.938.$$

Heptane Solution

$$\bar{\nu} = -0.008\,\mathrm{FT}_m + 39.232,$$
$$n = 5, \quad R = -0.935,$$
$$s = 0.073, \quad F = 20.700,$$

or

$$\bar{\nu} = -0.014\,\mathrm{FT}_e + 39.089,$$
$$n = 5, \quad R = -0.920,$$
$$s = 0.080, \quad F = 16.630.$$

Conclusions

This study has demonstrated the excellent correlation of the principal ultraviolet absorption maxima for compounds in several chemical classes with their corresponding unitary molecular structure indices. In multicyclic ring systems, the FT_m served also as a discriminator of topological structure groupings. In this series there was little dependence on the method of structure optimization (molecular mechanics or quantum chemical) in respect to the magnitude of the correlation coefficients. Also in this class, the FT_e, the FT_c, and the M_n provided satisfactory correlations in some of the FT_m-delineated groups.

In the series of conjugated dienes, the FT_m, FT_e, and M_n gave good correlations while FT_c and the electronic moment (M_n) were less satisfactory.

In the series of 24 substituted benzenes the FT_m, as a structure discriminator, showed that the halobenzenes may be treated as a separate class with excellent correlations based upon the FT_m, FT_e, and M_e indices. The remainder of the compounds in this series represented a wide variety of substituent classes and the correlation coefficients were less satisfactory. However, the general thrust of the results again suggested that correlations within classical chemical groups, with a sufficient number of examples in each, should be examined.

In a group of substituted monoalkybenezes, the FT_m and FT_e indices proved to be the most useful for spectra correlation in both the vapor phase and in heptane solution, with the FT_m being only marginally better than the FT_e, based on the F statistic of the models. But it was also interesting in this case that unsubstituted nitrobenzene was an outlier; it was not included in the basis for the equations recommended for absorption estimation.

For the purpose of estimating the absorption maximum for compounds similar to those considered herein, statistically robust equations were generated. For the aromatic hydrocarbons, the FT_m was employed although other indices would also serve as well, depending on which topological group was being considered. For the conjugated dienes, both FT_m and M_n may be used. In the substituted benzene series, FT_m, FT_e, and M_e provide equations with good statistics while, for more general substituents, the FT_m will serve to give preliminary estimates. In the case of the alkylnitrobenzenes, both FT_m and FT_e are quite satisfactory structure surrogates.

Future work in spectra-structure correlation should utilize the molecular similarity potential of the FT_m index and investigate the other indices for their performance in that capacity.

ACKNOWLEDGMENTS

The authors thank Dr. S. Edward Krikorian for helpful discussions and for providing several references. We thank Mr. Ronald J. Kassel and Dr. Herbert S. Aaron for assistance with compound nomenclature, and Ms. Ramona L. King for assistance with manuscript composition. We are extremely grateful to Mrs. Susan Molnar and Mrs. Mary F. King, whose continuing and unrelenting support made this work possible; and to the Master, who gave us the insight and who made the support possible.

References

1. L. J. Soltzberg and C. L. Wilkins, J. Am. Chem. Soc. **98**(13), 4066 (1976).
2. L. J. Soltzberg and C. L. Wilkins, J. Am. Chem. Soc. **99**, 439 (1977).
3. R. Wierl. Ann. Phys. (Liepzig) **8**, 521 (1931).
4. J. W. King and S. P. Molnar, *Dermal Transport and Molecular Structure: Direct Correlation via the Integrated Molecular Transform*, Presented at the 25th National Medicinal Chemistry Symposium, Ann Arbor, MI, 18–22 June 1996.
5. J. W. King and S. P. Molnar, *Structural Similarity Indices Generated from the Integrated Molecular Transform or the Normalized Molecular Moment*, Presented at the 2nd National Chemical Information Symposium, Charleston, SC, 14–18 July 1996.
6. J. W. King, R. J. Kassel, and B. B. King, Int. J. Quant. Chem., Quant. Biol. Symp. **17**, 27 (1990).
7. J. W. King and R. J. Kassel, Int. J. Quant. Chem., Quant. Biol. Symp. **18**, 289 (1991).
8. J. W. King and R. J. Kassel, Int. J. Quant. Chem., Quant. Biol. Symp. **19**, 179 (1992).
9. J. W. King, Int. J. Quant. Chem., Quant. Biol. Symp. **21**, 209 (1994).
10. S. P. Molnar and J. W. King, *Correlation of Mustard Analogs GC Retention Indices by the Integrated Molecular Transform*, Proceedings of the Ninth International Simulant Workshop, U.S. Army Chemical and Biological Defense Command (CBDCOM), Edgewood Area, Aberdeen Proving Ground, MD 21010-5423, 7–9 March 1995.
11. S. P. Molnar and J. W. King, Int. J. Quant. Chem., Quant. Biol. Symp. **22**, 201 (1995).
12. S. P. Molnar and J. W. King, Int. J. Quant. Chem., Quant. Biol. Symp. **23**, 1845 (1996).
13. J. W. King, Int. J. Quant. Chem., Quant. Biol. Symp. **20**, 139 (1993).
14. J. W. King and S. P. Molnar, J. Mol. Struct. (Theochem) **370**, 181 (1996).
15. J. W. King and S. P. Molnar, and Int. J. Quant. Chem., to appear.
16. G. R. Famini, R. J. Kassel, J. W. King, and L. Y. Wilson, Quant. Struct.-Act. Relat. **10**, 344 (1991).
17. R. M. Silverstein, G. C. Bassler, and T. C. Morrill, *Spectrometric Identification of Organic Compounds*, 5th ed. (Wiley, New York, 1991), pp. 298–304.
18. D. W. Brown, A. J. Floyd, and M. Sainsbury, *Organic Spectroscopy* (Wiley, New York, 1988), pp. 8–20.
19. R. D. Braun, *Introduction to Instrumental Analysis* (McGraw-Hill, New York, 1987), pp. 265–268.
20. ChemDraw Pro and Chem3D Pro are copyrighted (1995) products CambridgeSoft Corp., 875 Massachusetts Ave., Cambridge, MA 02139.
21. SYSTAT for Windows (Version 5, 1992), SYSTAT, Inc., 1800 Sherman Ave., Evanston, IL 60201-3793.
22. F. H. Clark, *Calculator Programming for Chemistry and the Life Sciences*, (Academic, New York, 1981), Chapter 4.
23. H. B. Klevins and J. R. Platt, J. Chem. Phys. **17**, 470 (1949).
24. H. B. Klevins, J. Chem. Phys. **18**, 1063 (1950).
25. H. H. Jaffe and M. Orchin, *Theory and Application of Ultraviolet Spectroscopy* (Wiley, New York, 1962).
26. H. B. Klevins and J. R. Platt, Survey of Vacuum Ultraviolet Spectra of Organic Compounds In Solution, Report of Task Order IX of Contract N6ORI, Office of Naval Research with the University of Chicago, in *Systematics of the Electronic Spectra of Conjugated Molecules, A Source Book* (Wiley, New York, 1964).

Dimerization of Dexanabinol by Hydrogen Bonding Accounts for Its Hydrophobic Character

EMIL POP, MARCUS E. BREWSTER
Pharmos Corporation, 2 Innovation Drive, Alachua, Florida 32615

Received 1 March 1997; revised 22 March 1997; accepted 23 April 1997

ABSTRACT: Dexanabinol, a dihydroxylated synthetic cannabinoid, is a member of the nonpsychotropic (+) 3S, 4S enantiomeric series. Experimental evidence suggests that dexanabinol might form aggregates (e.g., dimers) in which the two OH (a phenol and an allylic alcohol) groups are involved in hydrogen bonding. The extremely low solubility of dexanabinol in water implies that this interaction may not involve solvent molecules. A theoretical study of this phenomenon in the framework of the PM3 molecular approximation is described. Simple molecular models (phenol and 6-cyclohexene-1-methanol) were initially examined followed by extension of the calculations to dexanabinol. The results indicate that dimers of dexanabinol resulting from hydrogen bonding are more stable than the isolated molecules with the differences attributed to hydrogen bonding energies. It is suggested that the phenolic hydroxy group of one molecule forms a hydrogen bond with the allylic OH group of the second molecule and vice versa, resulting in dimers which contain two hydrogen bonds. The hydrogen bonds are more stable (6.14 kcal/mol) and the complex formed is more favored energetically when the phenol groups act as hydrogen bond donors and the allylic OH groups as acceptors. These interactions are also energetically more favored than those between dexanabinol and water (3.70 kcal/mol). The dexanabinol dimer manifested a lower dipole moment as compared to the monomer (1.211 vs. 2.221 debye) as well as a much larger log P (11.16 vs. 5.90), indicating strong hydrophobic character. The optimized structure shows that the OH groups involved in hydrogen bonds are oriented to the interior of the dimers, while the lipophilic side chains are oriented toward the exterior. These properties of the dimer may explain the low water solubility of dexanabinol.
© 1997 John Wiley & Sons, Inc. Int J Quant Chem **65**: 1057–1064, 1997

Correspondence to: E. Pop.

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061057-08

Introduction

Dexanabinol (HU-211), the (+) 3S, 4S-5′-(1′, 1′-dimethylheptyl)-Δ^6-7-hydroxy-tetrahydrocannabinol is a nonpsychotropic, synthetic cannabinoid currently in clinical development as a neuroprotective agent [1,2]. The extremely low solubility of this compound in water complicates its formulation as an intravenous drug. While the bulky lipophilic 1′,1′-dimethylheptyl side chain should reduce water solubility, the two hydroxyl groups present in the molecule of dexanabinol would be expected to induce a moderating influence. The compound was, however, found to be practically insoluble in water.

A possible explanation for this behavior is that the OH groups of dexanabinol cannot form hydrogen bonds with the solvent since they are already involved in more favorable intramolecular or intermolecular (dexanabinol-dexanabinol) interactions. Both OH groups present in dexanabinol (allylic at C-7 and phenol at C-3′) can participate in hydrogen bonding. While intramolecular hydrogen bonding are less probable due to the large distance between the two OH groups, intermolecular bonding is predictable [3–6] and can occur in various ways (i.e., phenol–phenol and allylic OH–allylic OH, phenol–allylic OH, etc.). Experimental evidence for this assumption is supported by the infrared (IR) spectra of dexanabinol. The shift of IR frequencies of the hydroxyl groups from 3590 to 3650 cm^{-1}, typical for free groups to lower values (3226 and 3424 cm^{-1}) is indicative of hydrogen bonding [7,8]. The relatively high melting point of dexanabinol (140–143°C) as compared to related compounds (e.g., the 6-methyl analog, dexanabinol pivalate, etc.) which are oils, is also suggestive of stabilization due to hydrogen bonding. Furthermore, acylation of the phenolic position of dexanabinol to form the acetate or other esters, paradoxically results in increased water solubility. The observation is consistent with incipient disruption of hydrogen bonds.

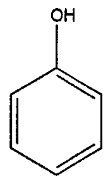
A theoretical study of this phenomenon has been performed. The PM3 molecular orbital approximation [9,10] was used for this purpose. The paradigm of the study included a comparison of the thermodynamic stability, as reflected by the calculated heat of formation (ΔHf), of dexanabinol monomer to the stability of a dimer resulting by double hydrogen bond formation of two molecules of dexanabinol. If the energy of dimers were found to be lower than the sum of two isolated monomers, the dimerization should be energetically favored. Moreover, the difference in energy should be a measure of the energy of the hydrogen bonding. Since the computations are rather complex with over 300 orbitals to be calculated, simple models (phenol and 6-cyclohexene-1-methanol) were used for initial studies. Vibrational spectra were only calculated for these models. Other physical properties relevant to solubility (dipoles, log P) are also evaluated.

Methods

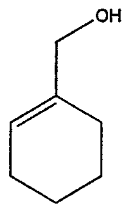
Theoretical studies were performed using PM3 molecular orbital approximation [9,10] which was included in the HyperChem (Hypercube, Inc., Waterloo, Ontario, Canada) version 5.0 software [11] run on a pentium Digital computer. PM3 uses a set of parameters derived from a larger number and variety of experimental versus calculated molecular properties, as compared to other semiempirical methods, including the AM1 procedure [12]. Typically, nonbonded interactions are less repulsive in the PM3 procedure [11]. Molecular models were constructed by the model builder of HyperChem. Geometry optimization was completed by using the Polak-Ribiere conjugate gradient algorithm method. The restricted Hartree–Fock (RHF) method was applied to the calculation of wave functions. Hydrogen bonds were displayed after molecule pairs were arranged so that the required conditions (hydrogen donor–acceptor distance less than 3.2 Å and the angle made by covalent bonds to the donor and acceptor atoms less than 120°) were fulfilled. Log P were calculated using a nonlinear regressional model in which all the descriptors used (molecular surface, volume, weight, etc.) are determined from the fully optimized structures [13], included in the QSAR package of the ChemPlus™ (version 1.5) extension for HyperChem.

Results and Discussion

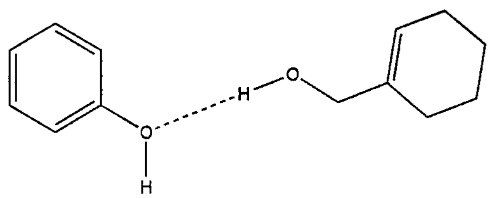
The hydrogen bonding in the selected model systems, i.e., phenol (**1**) and 6-cyclohexene-1-methanol (**2**) (Fig. 1), was considered in the initial



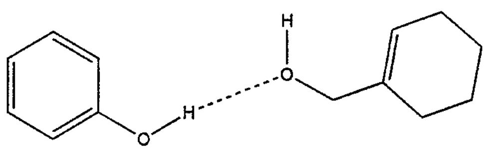
1



2



3



4

FIGURE 1. Structures of phenol (**1**), 1-cyclohexene-1-methanol (**2**), and two complexes (**3** and **4**) resulting by hydrogen bonding.

study, and heats of formation (ΔHf) determined for the optimized structures. Molecules **1** and **2** were then arranged so that the hydrogen bonding conditions were fulfilled. In model **3**, the allylic OH served as the hydrogen bond donor and the phenol as the acceptor, while in the model **4** the phenol is the hydrogen bond donor. Hydrogen bonds were displayed using the Hyperchem "show hydrogen bonds" and "recompute hydrogen bond" options. Both dimers were then reoptimized using the PM3 Hamiltonian. The calculated heats of formation of the complexes were compared to those of the monomers (Table I). In both cases, ΔHf of the complex resulting from hydrogen bonding was smaller than the sum of ΔHf's of **1** and **2**. It is reasonable to assume that these differences ($\Delta\Delta Hf$), calculated according to Eq. (1), represents the energies of the hydrogen bonding (the same results can be obtained by using the total binding energies of **1**, **2**, **3** and **4**, presented in Table I, instead of ΔHf's for these calculations). The results (Table I) indicate that the hydrogen bond energy was only 1.48 kcal/mol in the case of **3**, indicating the formation of a weak hydrogen bond, but 6.63 kcal/mol for **4** indicating a stronger hydrogen bond:

$$\Delta\Delta Hf = \Delta Hf_{(\mathrm{I})} + \Delta Hf_{(\mathrm{II})} - \Delta Hf_{(\mathrm{I+II})}, \quad (1)$$

where $\Delta Hf_{(\mathrm{I})}$ is the heat of formation of the phenol, $\Delta Hf_{(\mathrm{II})}$ the heat of formation of the 1-cyclohexen-1-methanol, and $\Delta Hf_{(\mathrm{I+II})}$, the heat of formation of the dimers (**3** or **4**).

Net charges and atomic orbital populations are presented in Table II. A higher degree of polarization of the positive hydrogen and negatively charged oxygen atoms can be noticed in the case of hydrogen-bonded molecules **3** and **4**. In the case of **3**, a strong participation of the Pz orbital of the allylic alcohol oxygen (hydrogen donor) was noted, while in the case of **4**, of the Pz orbital of the phenol oxygen (hydrogen donor) can be seen.

A vibrational analysis was performed and the vibrational (IR) spectra were calculated for the models. No negative frequencies appeared in the

TABLE I
Calculated binding energies (E), heats of formations (ΔHf), and estimated hydrogen bond energies ($\Delta\Delta Hf$) (kcal / mol)

Compound	E (kcal / mol)	ΔHf (kcal / mol)	($\Delta\Delta Hf$) (kcal / mol)
1	−1419.36	−21.85	—
2	−1931.91	−50.90	—
3	−3352.76	−74.23	1.48
4	−3357.90	−79.38	6.63
5	−6584.86	−154.05	—
6	−13171.30	−309.69	0.80
7	−13180.23	−320.69	6.14
8	−7026.68	−268.36	3.70

TABLE II
Net charges and atomic orbital electron populations for atoms involved in hydrogen bonding

Compound	Phenol O	Phenol H	Allylic alcohol O	Allylic alcohol H
	Net charge and atomic orbital electron population			
Phenol (**1**)	−0.228 Px: 1.275 Py: 1.243 Pz: 1.916	0.196 *s*: 0.803	—	—
Cyclohexene-1-methanol (**2**)	—	—	−0.306 Px: 1.574 Py: 1.227 Pz: 1.645	0.181 *s*: 0.819
Complex **3**	−0.248 Px: 1.286 Py: 1.422 Pz: 1.751	—	−0.335 Px:1.300 Pz: 1.302 Pz: 1.921	0.209 *s*: 0.791
Complex **4**	−0.271 Px: 1.660 Py: 1.418 Pz: 1.891	0.232 *s*: 0.678	−0.335 Px: 1.392 Py: 1.830 Pz: 1.297	—
Dexanabinol (**5**)	−0.231 Px: 1.361 Py: 1.333 Pz: 1.743	0.194 *s*: 0.805	−0.308 Px: 1.349 Py: 1.377 Pz: 1.775	0.184 *s*: 0.816
Dimer **6**	*a*: −0.252 Px: 1.545 Py: 1.692 Pz: 1.222 *b*:− 0.250 Px: 1.379 Py: 1.318 Pz: 1.759	*a*: 0.208 *s*: 0.792 *b*: 0.203 *s*: 0.797	*a*: −0.330 Px: 1.250 Py: 1.754 Pz: 1.514 *b*: −0.335 Px: 1.464 Py: 1.406 Pz: 1.653	*a*: 0.200 *s*: 0.800 *b*: 0.202 *s*: 0.798
Dimer **7**	*a*: −0.267 Px: 1.285 Py: 1.872 Pz: 1.315 *b*: −0.268 Px: 1.287 Py: 1.837 Pz: 1.350	*a*: 0.205 *s*: 0.794 *b*: 0.211 *s*: 0.788	*a*: −0.333 Px: 1.741 Py: 1.251 Pz: 1.535 *b*: −0.333 Px: 1.583 Py: 1.345 Pz: 1.599	*a*: 0.197 *s*: 0.803 *b*: 0.194 *s*: 0.805
Dexanabinol-water complex (**8**)	−0.264 Px: 1.378 Py: 1.336 Pz: 1.754	0.215 *s*: 0.784	−0.350 Px: 1.209 Py: 1.387 Pz: 1.944	0.197 *s*: 0.803

Note: *a* and *b* represent the two hydrogen bonds and atoms involved, respectively.

calculated IR spectra, indicating that valid minimum energy structures were obtained [11]. Recent results [14] indicated that out of PM3, AM1 and MNDO methods used for calculating IR frequency, PM3 showed the closest correspondence (which is generally about 10% too high in value of stretches) to experimental values. Indeed, calculated values of the O–H stretching vibrations were 3891 cm^{-1} for phenol (**1**) and 3812 cm^{-1} for the hydrogen bond containing complex **4**, indicating a batochromic shift as that observed experimentally for hydroxyl groups involved in hydrogen bonding.

The study of dimers of dexanabinol formed by hydrogen bonding was performed using the same rationale: the thermodynamic stability of the dexanabinol monomer (**5**) (Fig. 2), as reflected by calculated PM3 heat of formation (ΔHf) was first determined. The optimized geometry of **5** was then examined to verify the possibility of formation of intramolecular hydrogen bonds. The distance between the two oxygen atoms of the OH functionalities was too large (5.939 Å) to form hydrogen bondings. Two dexanabinol molecules were then arranged so that intermolecular hydro-

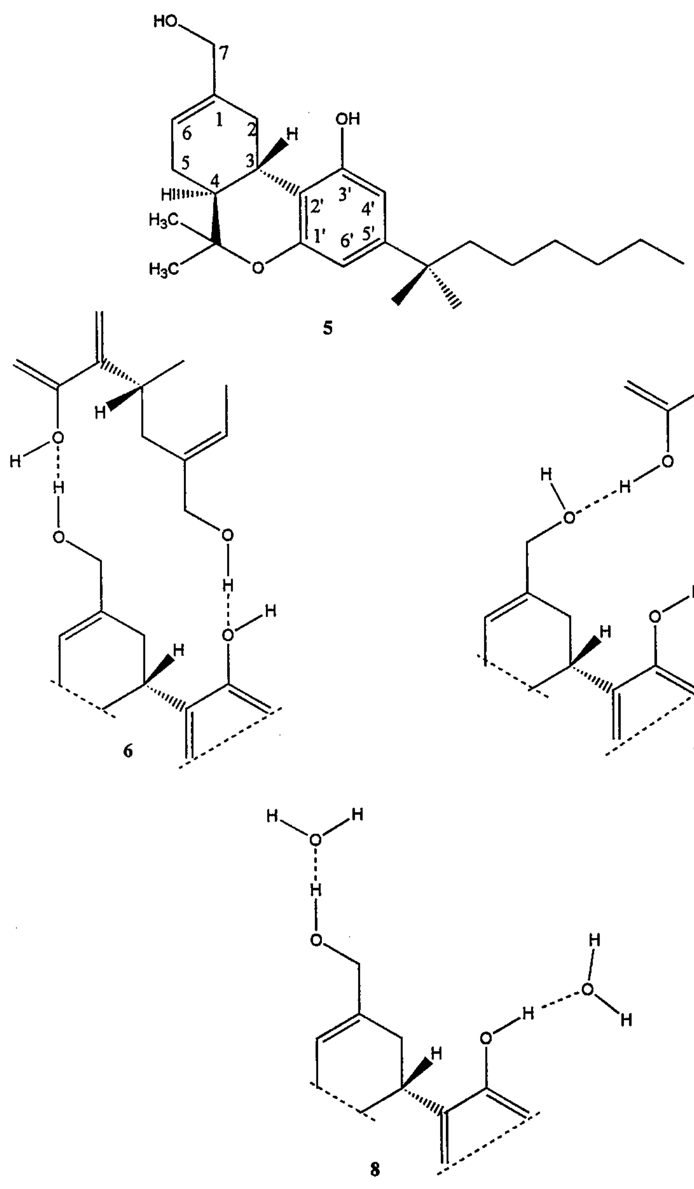


FIGURE 2. Structure of dexanabinol (**5**), two possible dimers (**6** and **7**), resulting by hydrogen bonding and the complex resulting by hydrogen bonding of **5** with water (**8**). Structures **6**–**8** are abbreviated.

gen bonding between OH groups could occur. Two models in which all four OH groups of the two dexanabinol molecules were involved in hydrogen bonding were built: one (models **6**) in which the allylic OH groups served as the hydrogen bond donors and the phenols as the acceptors, while in the other one (**7**) the phenols are the hydrogen donors and the allylic OH groups the acceptors. Due to sterical hindrance, models in which the phenol interacted with the juxtaposed phenol and allylic hydroxyl with the complementary allylic hydroxyl group could not be built. Also, models in which only one hydrogen bond was formed were not considered in this study. The geometries of the resulting dimers were then reoptimized using the PM3 procedure. Calculated heats of formations of dimers were compared with those of the monomers. The dimers had smaller ΔHf (-309.69 and -320.38 kcal/mol) than the sum of two monomers (-308.1 kcal/mol) (Table I), indicating that dimers were energetically favored. Equation (2) was used to calculate the hydrogen bond energies:

$$\Delta\Delta Hf = [2\Delta Hf_{(\mathrm{III})} - \Delta Hf_{(\mathrm{IV})}]/2 \qquad (2)$$

where $\Delta Hf_{(\mathrm{III})}$ is the heat of formation of dexanabinol and $\Delta Hf_{(\mathrm{IV})}$ is the heat of formation of the dimer **6** or **7**.

As in the case of the simpler systems discussed, the hydrogen bonds for the model in which the allylic OH was the hydrogen bond donor (dimer **6**) were less favored (0.795 kcal/mol) than those in which the phenol was the hydrogen donor (dimer **7**) (6.14 kcal/mol) (Table I). The polarization of the oxygen and hydrogen atoms involved in the bonding discussed above is also apparent in this case (Table II); i.e., the oxygen atoms have an increased negative charge, and the hydrogen atoms have an increased positive charge as compared to the monomer. In **6**, orbitals Py or the oxygen atoms have the larger contribution for one of the hydrogen bonds and the Px and Pz for the other, while in **7**, the Py orbitals were the major contributors to the highest occupied molecular orbitals (HOMO) for both oxygen atoms.

By following the same rationale, the hydrogen bonding of the dexanabinol monomer (**5**) with two molecules of water has been evaluated. The calculated (PM3) ΔHf for water is -53.46 kcal/mol. To both OH groups of optimized **5** a molecule of water was then added so that the hydrogen bond requirements were fulfilled. After the hydrogen bonds were built, the geometry of the trimolecular system, **8**, was reoptimized using the PM3 approximation. The calculated heat of formation of the supramolecular system was -268.36 kcal/mol. The energy of the hydrogen bonds ($\Delta\Delta Hf$) was calculated by using Eq. (3):

$$\Delta\Delta Hf = \left[(\Delta Hf_{(\mathrm{III})} + 2\Delta Hf_{\mathrm{H_2O}}) - \Delta Hf_{(\mathrm{V})}\right]/2, \quad (3)$$

where $\Delta Hf_{\mathrm{H_2O}}$ is the heat of formation of water and $\Delta Hf_{(\mathrm{V})}$ the heat of formation of **8**. The energy of the hydrogen bondings between dexanabinol and water was found to be 3.70 kcal/mol.

Due to the large volume of computations required, no vibrational spectra were calculated for these molecules.

These data suggest that hydrogen bonding stabilizes dexanabinol but that dimer formation was more energetic than interaction with water molecules. While the data presented above refer to the gas phase, thermodynamic stabilities including solvent effects are currently being examined. Obviously, the presence of water may significantly alter energetics of the system. It is possible that in spite of stronger hydrogen bonding associated with the dimer, dilution with concomitant increase in water concentration and decrease of dexanabinol concentration may weaken the dimeric interaction or even lead to dissociation. On the other hand, the observed poor solubility of dexanabinol in water, even at very low concentration is consistent with the presence of stable dimers that do not readily dissociate.

In this context it is of interest to examine other molecular properties of dexanabinol related to solubility. Dipoles are importance for solvation, and polar compounds generally manifest better solubility in polar solvents, such as water. The calculated dipole moment of the dexanabinol monomer (**5**) is 2.221 debye, while dimer **7** is quite symmetrical, having an even lower dipole moment (1.211 debye) (Table III). The dipole moment of the complex of **5** with two water molecules is higher (3.499 debye). Dipole moments indicate that the solubility of dexanabinol in water, especially in the form of the dimer, should be low. Table III includes data about the geometries of the hydrogen bonds such as the distances between the heavy atoms (oxygen) participating in the hydrogen bonds and the angles formed by O—H—O atoms. These parameters were not modified significantly during computations since the conditions were imposed a priori. The hydrogen is intermediate between two

TABLE III
Calculated geometries of hydrogen bonds, dipole moments and log *P*

Compound	O — H — O Hydrogen bond: O — O Distance (Å)	O — H — O Hydrogen bond: Angle (O — H — O) (degrees)	Dipole (debye)	log *P*
3	2.783	170.66	2.671	1.37
4	2.767	170.98	2.774	4.27
6	2.776	159.44	2.325	10.83
	2.813	159.69		
7	2.769	173.18	1.211	11.52
	2.770	174.87		
8	3.159	165.11	3.499	1.34
	2.763	171.41		

participating oxygen atoms, but not equidistant. By examining the shape of the dimer (Figs. 3 and 4), it can be seen that its structure is symmetrical, having the polar OH groups engaged in hydrogen bonding in the interior and the highly lipophilic dimethylheptyl side chains projected toward the exterior. These types of compounds are not expected to be soluble in water.

Log *P* (Table III), which are reliable indices for the lipophilicity of various compounds, indicate that while dexanabinol itself is lipophilic (log *P*: 5.90), the lipophilicity of dimer **7** is extremely high (log *P*: 11.52). Furthermore, we have begun to look at explicit interactions between the hydroxyl functions and water. To this end, the dihydrate **8** was examined as described. Addition of water molecules to the structure affects log *P* as indicated by a calculated value of 1.34 for the trimolecular system.

In summary, a dimer of dexanabinol resulting from hydrogen bonding is thermodynamically more favored than the individual molecules or hydrates of dexanabinol. The dimer is symmetrical as reflected by a low dipole moment and a high lipophilicity. Dimers appear to be stable and viable in the solid state (as suggested by experimental evidence such as high melting points and batochromic shifts of the OH in IR frequencies) and

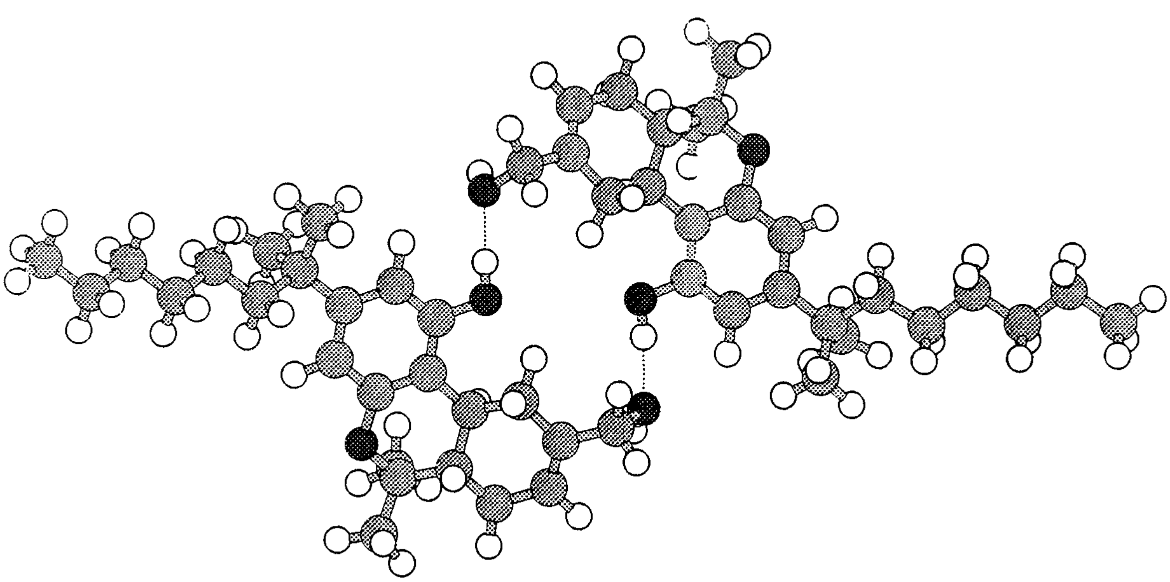


FIGURE 3. "Balls and cylinders" rendering of the optimized dimer **7**.

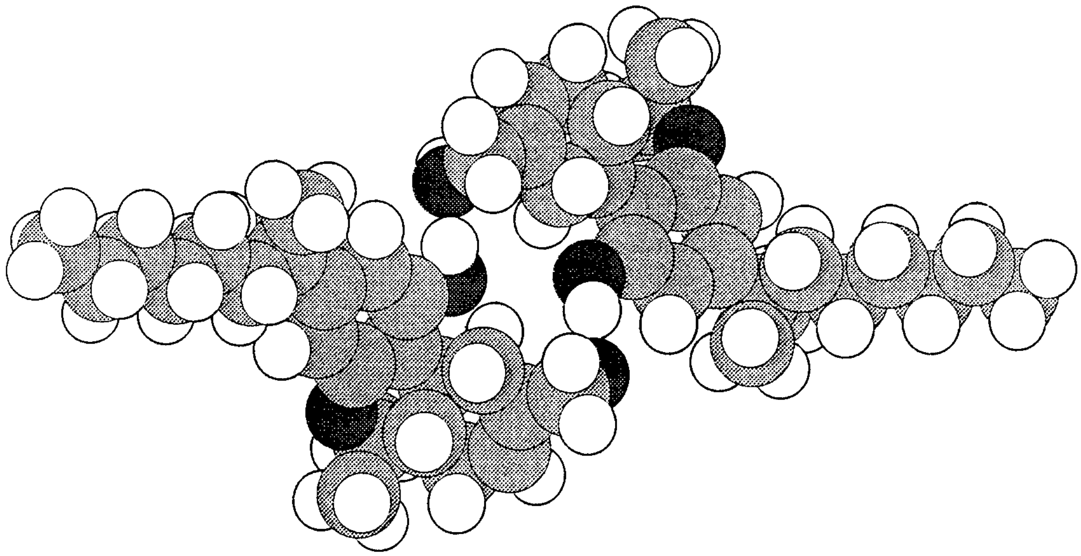


FIGURE 4. "Balls" representation of dimer **7**.

possible in gas phase as well. The poor solubility of dexanabinol in water indicate that dimers may not dissociate even when present at very low concentration. The theoretical findings presented above are supported by experimental evidence indicating poor aqueous solubility of this important drug candidate.

References

1. R. Mechoulam, N. Lander, A. Breuer, and Z. Zahalka, Tetrahedron-Asymmetry **1**, 315 (1990).
2. R. Mechoulam, W. A. Devane, and R. Glaser, in *Marijuana/Cannabinoids: Neurobiology and Neurophysiology*, L. Murphy and A. Batke, Eds. (CRC, Boca Raton, FL, 1992), pp. 1–33.
3. M. C. Etter, Acc. Chem. Res. **23**, 120 (1990).
4. Y. Ducharme and J. D. Wuest, J. Org. Chem. **53**, 5789 (1988).
5. S. K. Chang and A. D. Hamilton, J. Am. Chem. Soc. **110**, 1318 (1988).
6. R. Stadler, Prog. Colloid. Polym. Sci. **75**, 140 (1987).
7. M. Tichy, Adv. Org. Chem. **5**, 112 (1965).
8. M. R. C. Symons, Chem. Soc. Rev. **12**, 1 (1983).
9. J. J. P. Stewart, J. Comput. Chem. **10**, 209 (1989).
10. J. J. P. Stewart, J. Comput. Chem. **10**, 221 (1989).
11. HyperCharme 5.01. Hypercube, Inc. Waterloo, Ont., Canada, 1996.
12. M. J. S. Dewar and D. M. Storch, J. Am. Chem. Soc. **107**, 3893 (1985).
13. N. Bodor, Z. Gabanyi, and C.-K. Wong, J. Am. Chem. Soc. **111**, 3783 (1989).
14. D. M. Seeger, C. Korzeniewski, and W. Kowalchyk, J. Phys. Chem. **95**, 68 (1991).

On Characterization of Molecular Surfaces

MILAN RANDIĆ,[1] GORAN KRILOV[2]

[1] *Department of Mathematics and Computer Science, Drake University, Des Moines, Iowa 50311*
[2] *Department of Chemistry, Drake University, Des Moines, Iowa 50311*

Received 3 March 1997; revised 12 May 1997; accepted 12 May 1997

ABSTRACT: We consider the problem of quantitative characterization of the molecular surface. We start with a set of matrices, the elements of which give interatomic separation and higher powers of the separations. Averaged row sums of individual matrices suitably normalized give molecular profiles. The problem that we consider is how to generalize this approach to 2-dimensional and 3-dimensional objects. By using a large number of random points distributed over the molecular surface or molecular volume, respectively, we arrive at matrices from which one can extract invariants that offer a good characterization of the molecular surface and the molecular volume. It is suggested that the ratio V/S, where V and S are components of the volume and surface profile for a molecule, respectively, represents a novel shape index. © 1997 John Wiley & Sons, Inc. Int J Quant Chem **65**: 1065–1076, 1997

Introduction

Quantitative characterization of molecules remains one of the central subjects of mathematical chemistry, molecular modeling, the quantitative structure–activity relationship (QSAR), the computer manipulation of chemical structure, and chemical documentation. Most of the approaches of the past focused on atom–atom connectivity (when viewing a molecule as a molecular graph) [1] or molecular geometry (when considering molecules as 3-D objects) [2]. Such viewing of molecules corresponds to "ball-and-stick" model kits, which clearly portray bonding and spatial distributions of atoms. "Space-filling" model kits, on the other hand, better illustrate the molecular surface, crowded atoms, and protein cavities. Characterization of the molecular surface has received limited attention in the literature, in part, no doubt, due to the difficulties involved. As a result, we come across qualitative, rather than quantitative, attributes of the molecular surface. Mezey and collaborators [3] and Arteca [4], e.g., partitioned the molecular surface into regions of different curvature, which lead to regions of

Dedicated to Mrs. Per-Olov Löwdin, gracious companion of the Sanibel Symposia.

Correspondence to: M. Randić.

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061065-12

"catchment," bifurcations, saddle points, etc. The essential step is to associate a molecular configuration with a molecular surface whose shape features will characterize the configuration. In contrast, Bader and co-workers [5] considered the topography of contours of constant electron density and partitioned the molecular volume into regions associated with molecular fragments by considering the derivatives of the electron density function. While these are useful for the discussion of local molecular properties of the molecular surface or molecular volume, they do not give numerical characterization that can be associated with an individual molecular surface.

Part of the difficulty, as no doubt many have recognized, is that the characterization of the molecular surface is intimately connected with the characterization of the molecular shape. Molecular shape, in contrast to the molecular volume and molecular surface, cannot be represented by a single number. Shape is an elusive concept that has not yet been well defined and is the least well characterized one, despite the fact that we all have clear notions what is the shape of an object. Molecular shape is a critical factor for many molecular properties. For example, it is the dominant factor in olfactory sensations, as was recognized already by Ružička in the 1920s [6]. The similarity in shape, not the similarity in chemical structure, often determines the odorous quality of a molecule, as is reflected in the similarity of shapes of civetone, a macrocyclic compound, and two polycyclic steroids, which, as Prelog and Ružička have shown, all have a similar characteristic musk odor [7]. Molecular shape is important in circular dichroism and chirality. Enantiomers, if viewed in isolation, have identical all-physical properties, but when in an asymmetrical environment (either in the presence of polarized light of different orientation or in the proximity of an asymmetrical macromolecule), they show different behavior, because of the distinction between their shape and their mirror image shape, which becomes critical in a chiral environment.

Molecular Shape

Molecular shape is one of many common chemical concepts that remains elusive to quantification. Other concepts of great significance in chemistry that also are very common and whose characterization remains elusive or ambiguous include size [8], complexity,* branching [9], similarity [10], and structure itself [11]. In each case, we lack a clear definition or a generally accepted definition of the concept. Some may argue that these concepts ought to remain undefined, i.e., qualitative. We feel that Lord Kelvin was right when he said that [12]

> When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind. . . .

Some progress has been made toward quantification of the above-mentioned concepts* [9–11], including characterization of the molecular shape. A crude characterization of the molecular shape has been suggested for the interpretation of chromatographic retention indices of benzenoid polycyclic aromatic compounds [13]: the ratio L/B, where L is the length and B is the width of the smallest rectangle in which the structural formula of a benzenoid molecule can be inscribed. There are other geometrical indices of shapes that have been outlined in the literature [14, 15].

Kier [16] considered molecular graphs, which although devoid of strict geometrical information about a molecule, still allow one to discuss molecular shapes. Using as the extreme graph shapes the path graph, which corresponds to a linear chain structure, and the star graph (such as the molecular graph of neopentane), Kier arrived at the set of indices (called κ shape indices) that reflect to some extent molecular shape rather than molecular size as most topological indices do. An alternative approach would be to subtract from the topological index the index of size or consider a difference of two topological indices.† Despite the large number of available topological indices, some molecular properties are difficult to represent by a single dominant descriptor. In the case of alkanes, this is particularly true of critical temperature, critical pressure, and critical volume. It is of some interest to observe that the best single variable simple regression for critical temperature is ob-

* Some may think that size is a well understood concept but how is one to decide on the relative size of two molecules, one which has more atoms of smaller size and the other fewer atoms of larger size. Is molecular volume rather than the number of atoms to be a measure of size?

† This avenue has not yet been much investigated.

tained using the difference between the connectivity indices ${}^1\chi - {}^2\chi$ (the coefficient of regression $r = 0.889$, the standard error $s = 4.59°C$) [17]. Similarly, the descriptor ${}^2\chi - {}^3\chi$ gives the best single variable regression among 40 tested descriptors for surface tension in octanes [17]. Both cases suggest that the property considered, even for molecules of the same size (isomers of octane), depends on molecular shape, the characterization of which remains elusive. Obviously, we need more precise characterization of the molecular shape and molecular surface than were hitherto available.

Shape Profiles

Good *characterization* of a structure need not be unique but ought to be general enough so that it applies to structures not previously considered and even to structures not previously anticipated. This is different from the requirement for a good *representation* of a structure that has to be unique, so that the structure can be reconstructed from its *code*. Codes, however, imply rules for the numbering of atoms, ordering of attributes, etc. In contrast, the characterization of a molecule is based on structural *invariants* (i.e., mathematical properties of the structure) and, hence, is independent of representation, the numbering of atoms, or a selected pictographic form for the molecular skeleton. The problem is how to arrive at a general characterization of the molecular shape and molecular surface. This problem has been recently addressed and successfully resolved [18]. This does not mean that there is no room for alternative characterizations; it is very likely that they will emerge. However, the present approach satisfies the basic requirements of generality and is not computationally too involved. It has already been applied to a number of problems requiring characterization of molecules of different shape [19–22].

The approach is based on molecular geometry. We start with the distance matrix ${}^1D$ in which the matrix elements d_{ij} are given by the length of the distance between vertices i and j. In the case of hydrocarbons, the distance can be measured in units of CC bond length or, in other words, the element d_{ij} is given as the ratio of the interatomic distance and CC distance between vertices i and j. The row sum R_i of this matrix, when divided by the number of atoms in a molecule, gives the average distance from atom i to the rest of the atoms and represents a characterization of atom i. The average row sum R of this matrix, when divided by the number of atoms in a molecule, gives the average distance in the molecule, is a structural invariant, and represents a characterization of a molecule.

A single invariant is often not enough for the characterization of molecules. To arrive at additional descriptors, we consider next the distance matrix ${}^2D$ in which the matrix elements are given as $(d_{ij})^2$, where d_{ij} is the already defined distance between vertices i and j. The average row sum ${}^2R$ of this matrix divided by the number of atoms in a molecule gives an additional structural invariant of the molecule considered. One continues in this manner and generates the sequence

$$ {}^1R, {}^2R, {}^3R, {}^4R, {}^5R, {}^6R, {}^7R, \ldots . $$

Because the typical entry d_{ij} is larger than 1, raising matrix elements d_{ij} to a higher power increases the row sums that show typical divergence. To curb the divergence, we normalize the sequence by factorials, arriving at

$$ {}^1R, {}^2R/2!, {}^3R/3!, {}^4R/4!, {}^5R/5!, {}^6R/6!, {}^7R/7!, \ldots . $$

When the distance matrix involves all atoms in a molecule, we refer to the above sequences as the molecular profile. If the summation involves only atoms at the molecular boundary, we refer to them as shape profiles.

In Figure 1, we illustrate the profiles for two smaller benzenoids, chrysene and benzphenanthrene, reproduced from [18], which have the same number of Kekule valence structures and the same count of conjugated circuits [23], which reflects considerable similarity in their π-electron characteristics. However, they have different shape, and as we see from Figure 1, this is well reflected in their profiles. Clearly, chrysene, being elongated, has greater average interatomic separations. In the case considered, both the molecular (volume) profile and the shape (boundary) profile are the same, since all carbon atoms are at the molecular periphery. In large systems, in particular, in the case of peri-condensed benzenoids, the two will somewhat differ [18].

Extension of molecular profiles to true volume profiles, i.e., to the characterization of molecular volume, and, similarly, extension of the shape profiles to characterize the molecular surface are not straightforward. Such extensions require a conceptual "jump" from the characterization of discrete

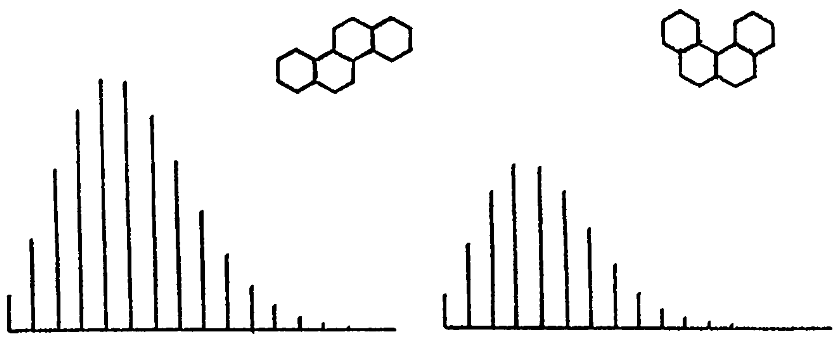


FIGURE 1. Molecular profiles for chrysene and benzphenanthrene.

objects, i.e., objects defined by a finite number of points in the space, to the characterization of a continuum. The bridge between the two was narrowed down with the introduction of bond profiles [20], a generalization of molecular profiles in which molecular connectivity is taken into account. This has been achieved by spacing "ghost" points along individual bonds in a molecule and averaging by taking into account also the presence of additional points in the structure. As the number of points in each bond is increased, one can bridge the gap between the discrete representation and continuum. It has been shown in the case of a cuboctahedron and a twist-cuboctahedron [21] that as n, the number of "ghost" points along each bond, increases the average bond profiles converge to a limit that represents a 1-D continuum. The problem that we will consider here is how to extend the same procedure to a 2-D continuum (and later the molecular surface) and a 3-D continuum (molecular volume).

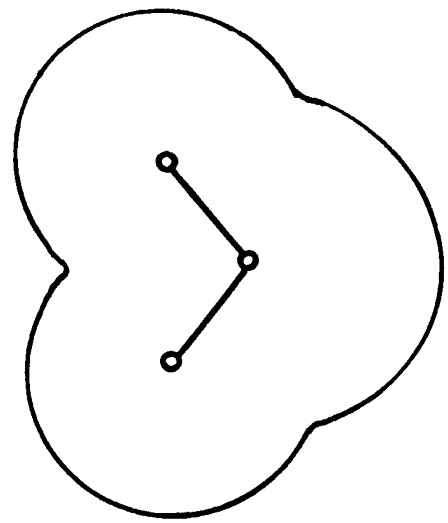


FIGURE 2. Van der Waals contour of H_2O molecule.

2-Dimensional Continuum

We will outline this generalization by considering a particular example: the van der Waals contour of the water molecule (Fig. 2). Clearly, the model of H_2O as a planar system is not adequate, but one can view this model as one projection of a 3-D water molecule. By using several such projections (along the lines of the work outlined by Jurs and collaborators [15]), one can arrive at a useful characterization of 3-D molecules. In the case of planar molecules, like most benzenoids, the approach presented here may suffice as an adequate characterization of molecules of difference shape.

What we need to do is to implant a large number of "ghost" points inside the molecular 2-D boundary and continue with calculating the average row sums as we did in the case of bond profiles [20, 21]. However, the points have to be uniformly distributed, which presents no problem in the case of bonds (particularly bonds of equal length). This would present insurmountable difficulties for irregular shapes, and even in the simplest cases (like a spherical shape, i.e., a circle in a 2-D projection), finding uniform distributions is impractical. The alternative is to place points at random, as in a typical Monte Carlo simulation. Not only does this make computations relatively simple, but a random distribution is also a better guarantee that we need not be concerned about systematic errors that may result from a regular placement of points.

In Table I, we list the computer profiles for a planar model of H_2O. They are based on van der

TABLE I
Profile for 2-D model of H_2O.

	1	2	3	4	5
500	0.398640 + 00	0.982942 − 01	0.183261 − 01	0.278948 − 02	0.361304 − 03
1000	0.403212 + 00	0.100323 − 00	0.188366 − 01	0.288239 − 02	0.374793 − 03
1500	0.398258 + 00	0.978132 − 01	0.181448 − 01	0.274585 − 02	0.353371 − 03
2000	0.396962 + 00	0.970965 − 01	0.179322 − 01	0.270121 − 02	0.345998 − 03
2500	0.397872 + 00	0.975155 − 01	0.180418 − 01	0.272201 − 02	0.349153 − 03
3000	0.397603 + 00	0.973593 − 01	0.179981 − 01	0.271362 − 02	0.347886 − 03
3500	0.396244 + 00	0.966925 − 01	0.1782-8 − 01	0.267995 − 02	0.342805 − 03
4000	0.395699 + 00	0.964036 − 01	0.177361 − 01	0.266226 − 02	0.339903 − 03
4500	0.396492 + 00	0.967967 − 01	0.178464 − 01	0.268445 − 02	0.343414 − 03
5000	0.396529 + 00	0.968175 − 01	0.178562 − 01	0.268751 − 02	0.344094 − 03
5500	0.397276 + 00	0.971776 − 01	0.179544 − 01	0.270688 − 02	0.347130 − 03
6000	0.397858 + 00	0.974466 − 01	0.180241 − 01	0.271997 − 02	0.349094 − 03
6500	0.397611 + 00	0.973311 − 01	0.179955 − 01	0.271493 − 02	0.348390 − 03
7000	0.397567 + 00	0.973060 − 01	0.179900 − 01	0.271436 − 02	0.348393 − 03
7500	0.397830 + 00	0.974200 − 01	0.180180 − 01	0.271941 − 02	0.349133 − 03
8000	0.398153 + 00	0.975688 − 01	0.180545 − 01	0.272562 − 02	0.349942 − 03
8500	0.398352 + 00	0.976677 − 01	0.180832 − 01	0.273168 − 02	0.350955 − 03
9000	0.398364 + 00	0.976648 − 01	0.180802 − 01	0.273070 − 02	0.350742 − 03
9500	0.398434 + 00	0.976951 − 01	0.180882 − 01	0.273226 − 02	0.350983 − 03
10,000	0.398096 + 00	0.975355 − 01	0.180474 − 01	0.272492 − 02	0.349951 − 03
10,500	0.398492 + 00	0.977236 − 01	0.180969 − 01	0.273427 − 02	0.351356 − 03
11,000	0.398810 + 00	0.978771 − 01	0.181384 − 01	0.274239 − 02	0.352625 − 03
11,500	0.398946 + 00	0.979429 − 01	0.181569 − 01	0.274619 − 02	0.353244 − 03
12,000	0.398950 + 00	0.979378 − 01	0.181534 − 01	0.274509 − 02	0.353007 − 03
12,500	0.399227 + 00	0.980663 − 01	0.181868 − 01	0.275134 − 02	0.353946 − 03
13,000	0.399804 + 00	0.983371 − 01	0.182576 − 01	0.276470 − 02	0.355958 − 03
13,500	0.399532 + 00	0.982048 − 01	0.182215 − 01	0.275758 − 02	0.354837 − 03
14,000	0.400188 + 00	0.985232 − 01	0.183069 − 01	0.277402 − 02	0.357345 − 03
14,500	0.400262 + 00	0.985532 − 01	0.183136 − 01	0.277507 − 02	0.357477 − 03
15,000	0.400186 + 00	0.985156 − 01	0.183030 − 01	0.277924 − 02	0.357129 − 03
15,500	0.400180 + 00	0.985120 − 01	0.183021 − 01	0.277279 − 02	0.357116 − 03
16,000	0.400043 + 00	0.984400 − 01	0.182818 − 01	0.276873 − 02	0.356476 − 03

Waals radii of 1.40 and 1.18 Å for oxygen and hydrogen atoms, respectively.

Volume Profiles

If the random points are distributed throughout the molecular interior defined by van der Waals atomic radii and then constructed from the corresponding matrices ^nD from which average $^nR/n!$ are extracted, we arrive at the molecular volume profile of a molecule. In Table II, we give the results for the first 10 powers of the separations using from 500 points to 12,500. As we see from Table II, there are variations between the individual calculations, as there should be for a Monte Carlo simulation, but these are not excessive, of the order of 2%, thus affecting the third digit. If higher accuracy is warranted, one should continue to increase the number of points or, what amounts to the same from a statistical point of view, one should start afresh and repeat the calculations again and again. To increase the accuracy by the factor of 10, one should increase the number of points by the factor of 100. Hence, the individual computations reported here in steps of 500 random points have to be extended to 50,000 points (four times more than we choose to carry out) in order to visibly improve the accuracy (by fixing the next digit in the individual profile components).

Table II shows the cumulative effect of Monte Carlo simulation which, if continued, would show

TABLE II
Profile for 3-D model of H_20.

	1	2	3	4	5
500	0.365020 + 00	0.825372 − 01	0.141649 − 01	0.199341 − 02	0.239725 − 03
1000	0.360752 + 00	0.806397 − 01	0.136990 − 01	0.191022 − 02	0.227781 − 03
1500	0.359130 + 00	0.798093 − 01	0.134674 − 01	0.186407 − 02	0.220512 − 03
2000	0.358998 + 00	0.797402 − 01	0.134524 − 01	0.186187 − 02	0.220249 − 03
2500	0.358495 + 00	0.795053 − 01	0.133968 − 01	0.185277 − 02	0.219100 − 03
3000	0.359606 + 00	0.799771 − 01	0.135154 − 01	0.187590 − 02	0.222420 − 03
3500	0.359646 + 00	0.800051 − 01	0.135219 − 01	0.187560 − 02	0.222426 − 03
4000	0.359939 + 00	0.801490 − 01	0.135607 − 01	0.188321 − 02	0.223628 − 03
4500	0.360070 + 00	0.801950 − 01	0.135694 − 01	0.188424 − 02	0.223701 − 03
5000	0.359460 + 00	0.799247 − 01	0.135041 − 01	0.187298 − 02	0.222159 − 03
5500	0.359982 + 00	0.801470 − 01	0.135570 − 01	0.188201 − 02	0.223380 − 03
6000	0.358818 + 00	0.796593 − 01	0.134435 − 01	0.186300 − 02	0.220839 − 03
6500	0.358502 + 00	0.795258 − 01	0.134140 − 01	0.185846 − 02	0.220302 − 03
7000	0.359090 + 00	0.797519 − 01	0.134630 − 01	0.186616 − 02	0.221270 − 03
7500	0.359712 + 00	0.800149 − 01	0.135264 − 01	0.187737 − 02	0.222864 − 03
8000	0.359283 + 00	0.798336 − 01	0.134831 − 01	0.186982 − 02	0.221807 − 03
8500	0.359194 + 00	0.797893 − 01	0.134709 − 01	0.186742 − 02	0.221432 − 03
9000	0.359205 + 00	0.797873 − 01	0.134692 − 01	0.186690 − 02	0.221319 − 03
9500	0.359247 + 00	0.797956 − 01	0.134697 − 01	0.186677 − 02	0.221280 − 03
10,000	0.359758 + 00	0.800140 − 01	0.135219 − 01	0.1875806 − 02	0.222526 − 03
10,500	0.359155 + 00	0.797606 − 01	0.134626 − 01	0.186579 − 02	0.221179 − 03
11,000	0.358664 + 00	0.795377 − 01	0.134052 − 01	0.185503 − 02	0.219565 − 03
11,500	0.359039 + 00	0.796976 − 01	0.134434 − 01	0.186163 − 02	0.220476 − 03
12,000	0.358981 + 00	0.796720 − 01	0.134362 − 01	0.186010 − 02	0.220212 − 03
12,500	0.359006 + 00	0.796870 − 01	0.134405 − 01	0.186093 − 02	0.220335 − 03

	6	7	8	9	10
500	0.252854 − 04	0.238194 − 05	0.203101 − 06	0.158384 − 07	0.113898 − 08
1000	0.238349 − 04	0.222835 − 05	0.188636 − 06	0.146091 − 07	0.104369 − 08
1500	0.228807 − 04	0.212049 − 05	0.177897 − 06	0.136522 − 07	0.966411 − 09
2000	0.228518 − 04	0.211735 − 05	0.177560 − 06	0.136173 − 07	0.963064 − 09
2500	0.227336 − 04	0.210718 − 05	0.176824 − 06	0.135735 − 07	0.961099 − 09
3000	0.231527 − 04	0.215295 − 05	0.181238 − 06	0.139552 − 07	0.991039 − 09
3500	0.231395 − 04	0.214998 − 05	0.180804 − 06	0.139048 − 07	0.986067 − 09
4000	0.233012 − 04	0.216899 − 05	0.182794 − 06	0.140925 − 07	0.100219 − 08
4500	0.233013 − 04	0.216817 − 05	0.182648 − 06	0.140752 − 07	0.100053 − 08
5000	0.231238 − 04	0.215039 − 05	0.181061 − 06	0.139469 − 07	0.990999 − 09
5500	0.232614 − 04	0.216371 − 05	0.182191 − 06	0.140321 − 07	0.996749 − 09
6000	0.229756 − 04	0.213577 − 05	0.179763 − 06	0.138416 − 07	0.100219 − 08
6500	0.229241 − 04	0.213166 − 05	0.179490 − 06	0.138269 − 07	0.982541 − 09
7000	0.230267 − 04	0.214111 − 05	0.180263 − 06	0.138837 − 07	0.986331 − 09
7500	0.232184 − 04	0.216121 − 05	0.182140 − 06	0.140418 − 07	0.998488 − 09
8000	0.230930 − 04	0.214822 − 05	0.180938 − 06	0.139413 − 07	0.990784 − 08
8500	0.230440 − 04	0.214269 − 05	0.180387 − 06	0.138922 − 07	0.986814 − 09
9000	0.230249 − 04	0.214004 − 05	0.180073 − 06	0.138597 − 07	0.983845 − 09
9500	0.230181 − 04	0.213912 − 05	0.179970 − 06	0.138495 − 07	0.982948 − 09
10,000	0.231626 − 04	0.215367 − 05	0.181270 − 06	0.139545 − 07	0.990676 − 09
10,500	0.230106 − 04	0.213879 − 05	0.179979 − 06	0.138534 − 07	0.983465 − 09
11,000	0.228072 − 04	0.211659 − 05	0.177833 − 06	0.136670 − 07	0.968740 − 09
11,500	0.229131 − 04	0.212728 − 05	0.178794 − 06	0.137450 − 07	0.974523 − 09
12,000	0.228748 − 04	0.212257 − 05	0.178286 − 06	0.136966 − 07	0.970389 − 09
12,500	0.228898 − 04	0.212413 − 05	0.178432 − 06	0.137088 − 07	0.971323 − 09

lesser and lesser oscillations in subsequent rows of the Table II. To determine that the accuracy of computations increased with an increase of n, the number of random points used, the entries in Table II should be compared with those in Table III, in which each row shows the same profile components, but each time for a different set of 500 random points.

Surface Profiles

To obtain a characterization of the molecular surface, we restrict random points to the surface and discard all points that correspond to the molecular interior. We could first generate, using pseudorandom numbers, the x, y coordinates of a point in the molecular interior using the same geometrical restrictions as used when considering molecular volume profile. However, instead of obtaining the z-coordinate by selecting a third random number, one should calculate the z-coordinate so that it is on the molecular surface. The coordinate system could be oriented so that half of the molecular van der Waals surface of water is above the x, y plane ($z > 0$) and half is below. One selects one or the other half of the surface depending on whether the random number is even or odd. (Other choices can be considered, such as alternation of the two halves of the surface.) Such an approach appears suitable when considering arbitrary shapes. However, since we have represented each atom by a sphere (having van der Waals radius for the atom considered), for our approach, it is better to use polar coordinates rather than Cartesian coordinates. The following are the steps that we performed in order to obtain uniformly distributed random points over the molecular surface.

TABLE III
Variations in the computed 3-D profile based each time on different 500 random points.

	1
500	0.365020 + 00
500	0.356483 + 00
500	0.355885 + 00
500	0.358604 + 00
500	0.356482 + 00
500	0.365165 + 00
500	0.359886 + 00
500	0.361990 + 00
500	0.361113 + 00
500	0.353973 + 00

1. Each atom is assigned a percent surface area. The obtained percentages delineate individual atoms, and when the first random number is generated, this leads to the selection of one of the atoms in a molecule. The corresponding probabilities are given as $P_n = R_n^2/(\Sigma R_i^2)$, where R_i are the van der Waals radii or atoms in a molecule.
2. The next two random numbers determine the magnitude of the Euler angles ϕ and ϑ (in intervals $[0, 2\pi]$ and $[0, \pi]$, respectively). By knowing the angles ϕ and ϑ, we calculate the Cartesian coordinates of the point on the surface of a sphere using

$$x_n = r_n \cos\phi \sin\vartheta$$
$$y_n = r_n \sin\phi \sin\vartheta$$
$$z_n = r_n \cos\vartheta .$$

3. Having a point on one of the atomic spheres, we first find other spheres that overlap with the one considered. For example, if the random number selects one of the hydrogen atoms of H_2O, then only the oxygen atom overlaps the hydrogen sphere, while the other hydrogen sphere is disjoint.
4. Next, we check if the obtained point (x_n, y_n, z_n) is inside any of other overlapping spheres. If it is, the point is discarded since it belongs to the overlapping region; if not, it is added to the list of random points representing the molecular surface.

When considering the molecular volume, the approach is modified by selecting an additional random number which determined the r-coordinate of a point (in the domain $0 \le r \le r_n$) so that it is inside the atomic sphere.

In Table IV, we list the derived molecular surface profiles for H_2O using the first 10 powers of the profile matrices. For the same n, we obtain a denser distribution of the points on surface than was the case with the density of the points when

TABLE IV
H_2O molecule surface profile.

	1	2	3	4	5
500	0.510572 + 00	0.149417 + 00	0.313542 − 01	0.517469 − 02	0.707465 − 03
1000	0.512496 + 00	0.150115 + 00	0.315247 − 01	0.520770 − 02	0.707465 − 03
1500	0.512868 + 00	0.150280 + 00	0.315743 − 01	0.521905 − 02	0.714813 − 03
2000	0.512192 + 00	0.149838 + 00	0.314310 − 01	0.518727 − 02	0.709403 − 03
2500	0.511507 + 00	0.149411 + 00	0.312797 − 01	0.514962 − 02	0.702193 − 03
3000	0.511120 + 00	0.149214 + 00	0.312281 − 01	0.514020 − 02	0.700874 − 03
3500	0.510536 + 00	0.148884 + 00	0.311281 − 01	0.511884 − 02	0.697292 − 03
4000	0.510601 + 00	0.148870 + 00	0.311173 − 01	0.511576 − 02	0.696695 − 03
4500	0.511103 + 00	0.149135 + 00	0.311957 − 01	0.511324 − 02	0.699474 − 03
5000	0.511314 + 00	0.149278 + 00	0.312449 − 01	0.514388 − 02	0.701522 − 03
	6	**7**	**8**	**9**	**10**
500	0.828731 − 04	0.851881 − 05	0.782251 − 06	0.650601 − 07	0.495514 − 08
1000	0.835794 − 04	0.859940 − 05	0.790163 − 06	0.657337 − 07	0.500501 − 08
1500	0.838940 − 04	0.863988 − 05	0.794691 − 06	0.661822 − 07	0.504492 − 08
2000	0.831424 − 04	0.855130 − 05	0.785593 − 06	0.653513 − 07	0.497638 − 08
2500	0.820215 − 04	0.840456 − 05	0.768981 − 06	0.636924 − 07	0.482798 − 08
3000	0.818738 − 04	0.839112 − 05	0.768003 − 06	0.636399 − 07	0.482671 − 08
3500	0.813750 − 04	0.833141 − 05	0.761713 − 06	0.630470 − 07	0.477604 − 08
4000	0.812838 − 04	0.831970 − 05	0.760406 − 06	0.629172 − 07	0.476441 − 08
4500	0.816691 − 04	0.836553 − 05	0.765189 − 06	0.633627 − 07	0.480192 − 08
5000	0.819656 − 04	0.840181 − 05	0.769038 − 06	0.637239 − 07	0.483234 − 08

calculating the molecular volume. Hence, we expect smaller variations in the computed molecular profiles for the same n. As we see from Table IV, we have convergence in the third digit as we approach 5000 random points.

Discussion

Although we illustrated the approach on one of the simplest molecules, H_2O, nevertheless, it is clear from the exposition that the approach is general. All that is required are the coordinates of the centers for atoms in a molecule and the van der Waals radii of all atoms. The interpretation of the results is also self-evident: The volume profile gives a characterization of the molecular volume, while when the random points are restricted to the molecular surface, we obtain a molecular surface profile, a set of numbers that represent a characterization of the molecular surface. If one applies the same procedure to 2-D objects, the volume profiles will correspond to area profiles and the surface profile will correspond to contour profiles. Hence, we can, within the same framework, consider the characterization of contours representing equal density, or equal potential, whether these are represented in a plane or in 3-D space.

Do the components in the derived volume and surface profiles have other structural interpretations? The average point-to-point distance and the averages of the corresponding powers of point-to-point distances represent a clear mathematical explanation for the numerical values derived. One should also not forget that the process of averaging the row contribution of the matrix is associated with a loss of information. Hence, there is no guarantee that the derived profiles will be unique or that from a given profile one could reconstruct the structure. However, as has been illustrated with the characterizations of molecules based on graph theoretical invariants (topological indices), although complete reconstruction need not be feasible, one can often narrow the number of structures that can have given values for selected topological indices [23–26]. We can see already from Figure 1 that if the profile that corresponds to chrysene is given (without our knowledge to which benzenoid it belongs) that we can immediately eliminate benzphenanthrene as a possible candi-

date, since the magnitudes of its profile components are too small to correspond to the data considered. In a similar fashion, the problem of generation and reconstruction of structures with a specified volume profile or surface profile can be expected to yield a short list of candidate structures.

Novel Shape Index

The volume and surface profiles when combined may offer additional insight on the structure considered. If a molecule is spherical, the ratio of the volume-to-surface (V/S) is simply $r/3$, where r is the radius of the sphere. However, if the object is cylindrical, such as obtained by inserting cylindrical parts between the two halves of a sphere (see Fig. 3), the ratio V/S will vary and also will be a function of L, the length of the cylindrical parts. For example, if we assume that $L = r$, $L = 2r$, and $L = 3r$, we obtain for the ratio V/S $0.388889r$, $0.4166667r$, and $0.4333333r$, respectively. Clearly, V/S is a function of the shape (here, simply a function of the length of the elongated structure).

We can introduce V/S as an index of molecular shape by using computed molecular profiles. In Table V, we list for the computed volume and surface profiles of H_2O the corresponding ratios V/S. Again, as n, the number of points, increases, the oscillatory behavior of the computed ratio V/S should attenuate, although in view that both the numerator and the denominator are based on random points, the convergence of the ratio V/S is going to be somewhat slower. Based on 5000 points, we obtain for V/S, when the powers used are $k = 1$–10, 0.6999, 0.5906, 0.4322, 0.3641, 0.3167, 0.2821, 0.2559, 0.2354, 0.2189, and 0.2051. Although both the volume and the surface of the van der Waals model of H_2O decrease quickly with the increased powers of k (see Tables II and Table IV, respectively), the quotient V/S tapers off rather slowly. Thus, while V decreased by 10^8 when going from $k = 1$ to $k = 10$, the ratio has decreased by a third. This suggests that the new shape descriptors contain a considerable amount of structural information, as they should, in view of the immense variability of shapes for molecules of the same volume and the same surface.

Open Problems

We have outlined a general procedure for the construction of a structurally related ordered set of descriptors for molecules viewed as imbedded in 3-D space and descriptors for the molecular surface of such molecules. The present work can be viewed as a seminal article, even though it has

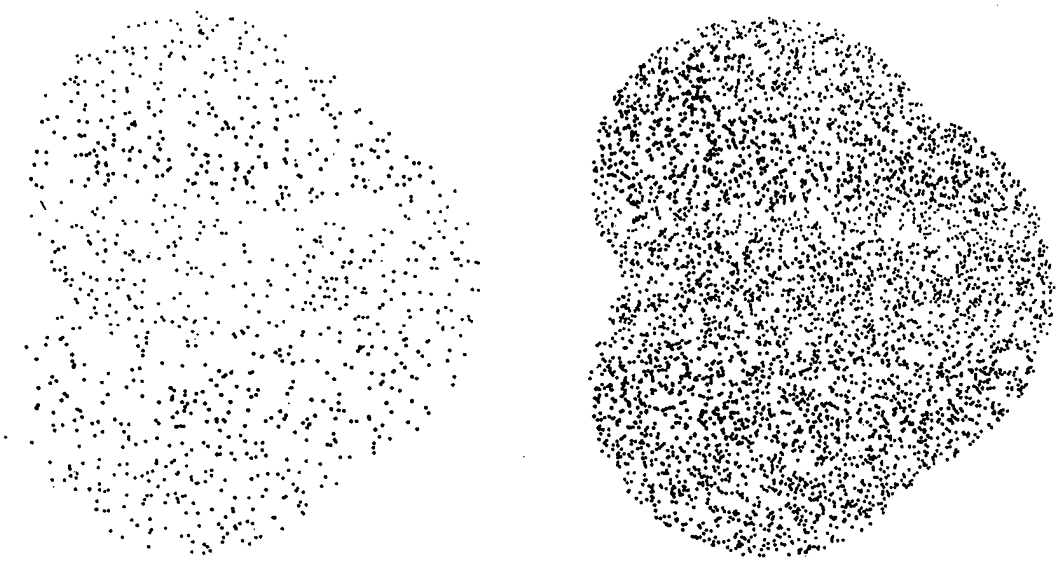


FIGURE 3. Representation of van der Waals contour of H_2O molecule by 1000 and 5000 random points, respectively.

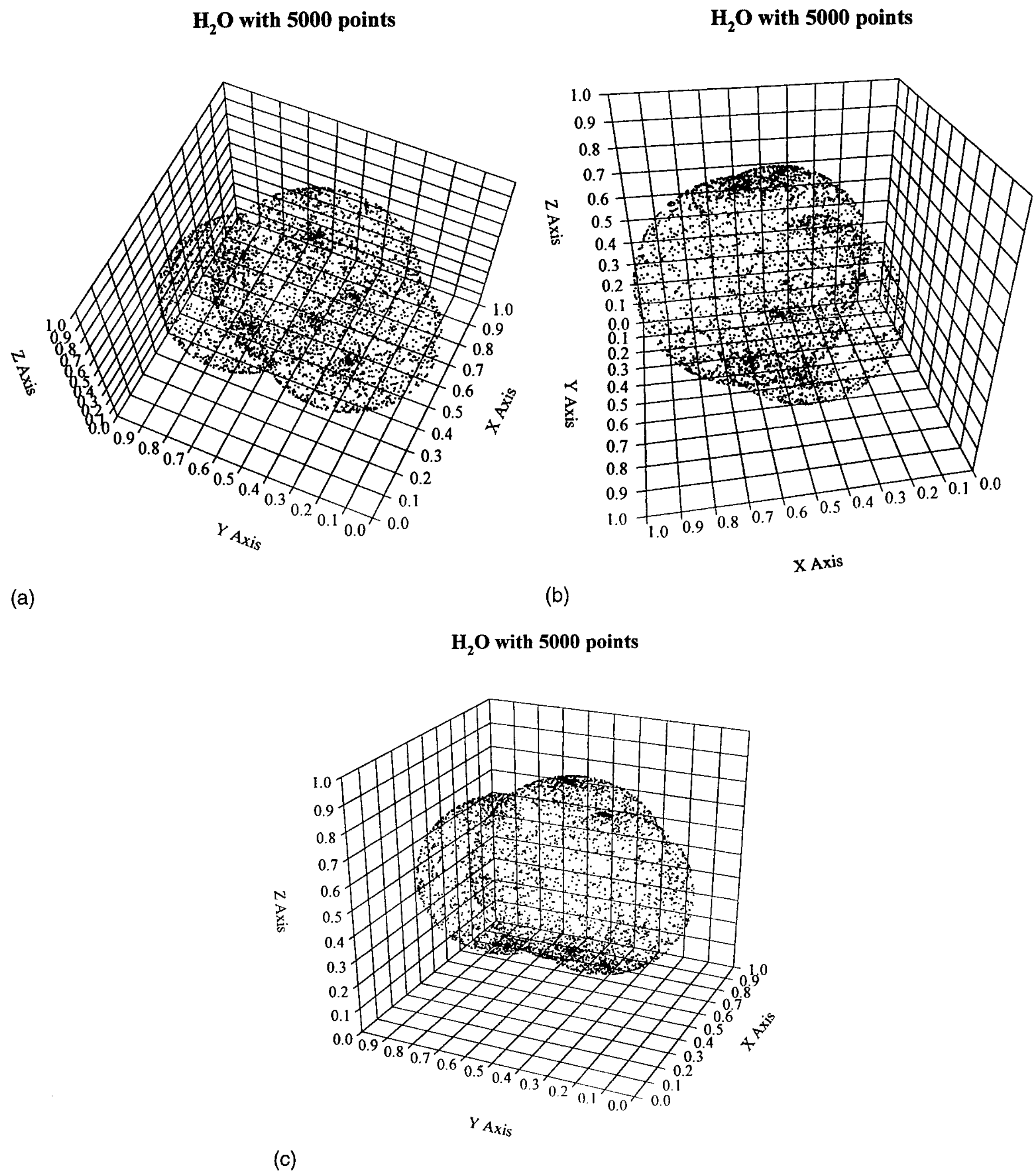


FIGURE 4. Three different views of 3-D model of H_2O molecule based on van der Waals radii for oxygen and hydrogen atoms derived by using 5000 random points.

evolved from profiles derived when atoms are represented as vertices of molecular graphs or points in space. The attribute "general" is indeed appropriate because not only are there no restrictions on the size and conformations of molecules considered, but there are also no restrictions on further modifications of the approach when instead of molecular volume and molecular surface other spatial functions of molecular electron density are considered, including the electron density itself. So, on the one hand, random points can be weighted to simulate electron density and, on the other hand, points can be distributed along computed equipotential surfaces. Equally, one can fo-

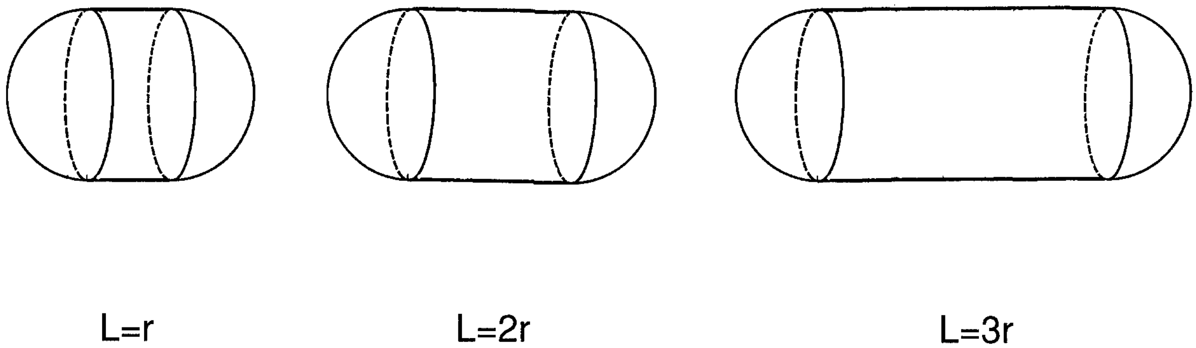


FIGURE 5. Cylinders of different length capped by hemispheres representing objects of similar but different shapes.

cus attention only on a local molecular environment and derive local molecular descriptors by suitable partitioning of the molecular volume and surface.

There are a number of open problems, technical or not, that deserve attention. It would be nice to establish a preferred number of random points per atom which would suffice in typical structure–activity studies. As we see from this work, about 10^3 random points define the components of the profiles to about 1 part in 10^3. This, of course, will suffice for most data reduction problems where experimental accuracy is at the level of 1%, but, occasionally, higher precision may be required. In contrast to some quantum chemical computations, we are not (yet) approaching hardware computational restrictions, since the computations increase with N, the number of random points, approximately linearly.

TABLE V
The quotient V/S.

	1	2	3	4	5
500	0.7126	0.5524	0.4518	0.3852	0.3389
1000	0.7033	0.5372	0.4345	0.3668	0.3196
1500	0.7013	0.5311	0.4269	0.3572	0.3085
2000	0.7002	0.5322	0.4280	0.3589	0.3105
2500	0.6780	0.5322	0.4283	0.3598	0.3120
3000	0.7018	0.5360	0.4328	0.3648	0.3173
3500	0.7016	0.5374	0.4344	0.3664	0.3190
4000	0.7020	0.5384	0.4358	0.3681	0.3210
4500	0.7015	0.5377	0.4350	0.3671	0.3198
5000	0.6999	0.5906	0.4322	0.3641	0.3167
	6	**7**	**8**	**9**	**10**
500	0.3051	0.2796	0.2596	0.2434	0.2299
1000	0.2852	0.2591	0.2387	0.2222	0.2085
1500	0.2727	0.2454	0.2239	0.2063	0.1916
2000	0.2749	0.2476	0.2260	0.2084	0.1935
2500	0.2772	0.2507	0.2299	0.2131	0.1991
3000	0.2828	0.2566	0.2360	0.2193	0.2053
3500	0.2844	0.2581	0.2374	0.2205	0.2065
4000	0.2867	0.2607	0.2404	0.2240	0.2103
4500	0.2853	0.2592	0.2387	0.2221	0.2084
5000	0.2821	0.2559	0.2354	0.2189	0.2051

Of immediate interest is to investigate the sensitivity of the computed profiles to geometrical variations of flexible molecules and variations between profiles of conformers, rotational or positional. Another topic of immediate interest concerns the application of the approach to structure–property–activity studies. These are, no doubt, to follow soon, because our model for the first time allows researchers to quantify the visual impressions of molecular models observed either on computer screens in virtual space–time or when "playing" with them in the real space–time, and these molecular models can be models where a molecule is represented by "ball-and-stick" as well as by space-filling models, both of which are often combined in current molecular computer graphics. The former correspond to the atom-point representation, and the second, to the van der Waals representation of atoms. But, also, other molecular models can be considered with the present scheme. For example, it should be possible to quantify, i.e., "translate," an image into a numerical characterization, such as those of the strands of β-sheets or α-helices in proteins. Moreover, one can expect that in the not so distant future software may appear which will "read" stereoviews of molecules (including complex structures like proteins) and produce their profile characterizations, either as "bond" profiles or space-filled profiles. This task is beyond our interest (i.e., beyond our expertise)

and we do not even wish to speculate whether it presents a challenge or not for an experienced computer software expert. But those looking for a challenge may consider another problem that has barely been "touched" in mathematical chemistry: The inverse structure problem. The question is: How and to what extent one can reconstruct a structure given the collection of its invariants? The pioneering work in this direction has just received attention in recent years by considering the reconstruction of molecular graphs from several well-known topological indices [23–26].

References

1. N. Trinajstić, *Chemical Graph Theory* (CRC Press, Boca Raton, FL, 1992).
2. A. T. Balaban, Ed., *From Chemical Topology to Three-Dimensional Geometry* (Plenum Press, New York, 1997).
3. P. G. Mezey, *Shape in Chemistry: An Introduction to Molecular Shape and Topology* (VCH, New York, 1993). P. G. Mezey, J. Math. Chem. **2**, 299 (1988). G. A. Arteca and P. G. Mezey, Int. J. Quantum Chem., Quantum Biol. Symp. **14**, 133 (1987). G. A. Arteca and P. G. Mezey, Int. J. Quantum Chem., Quantum Biol. Symp. **15**, 33 (1988). G. A. Arteca and P. G. Mezey, Int. J. Quantum Chem., Quantum Chem. Symp. **23**, 305 (1989). G. A. Arteca, G. A. Heal, and P. G. Mezey, Theor. Chim. Acta **76**, 377 (1990).
4. G. A. Arteca, Int. J. Quantum, Chem., Quantum Chem. Symp. **27**, 547 (1993). G. A. Arteca, J. Phys. Chem. **97**, 13831 (1993). G. A. Arteca, Biopolymers **33**, 1829 (1993). G. A. Arteca, J. Comput. Chem. **15**, 633 (1994). G. A. Arteca, Phys. Rev. E **49**, 2417 (1994). G. A. Arteca, Adv. Comput. Biol. **1**, 1 (1994). G. A. Arteca, Can. J. Chem. **73**, 241 (1995).
5. R. F. W. Bader, S. G. Anderson, and A. J. Duke, J. Am. Chem. Soc. **101**, 1389 (1979). R. F. W. Bader, T. T. Nguyen-Dang, and Y. Tal, J. Chem. Phys. **70**, 4316 (1979). R. F. W. Bader and T. T. Nguyen-Dang, Adv. Quantum Chem. **14**, 63 (1981). Y. Tal, R. F. W. Bader, T. T. Nguyen-Dang, M. Ohja, and S. G. Anderson, J. Chem. Phys. **74**, 5162 (1981).
6. L. Ružička, Chem. Zeit. **44**, 93, 129 (1920).
7. V. Prelog and L. Ružička, Helv. Chim. Acta **27**, 66 (1944).
8. S. H. Bertz, J. Am. Chem. Soc. **103**, 3599 (1981). S. H. Bertz, J. Chem. Soc. Chem. Commun. 818 (1981). S. H. Bertz, J. Am. Chem. Soc. **104**, 5801 (1982). S. H. Bertz, Discr. Appl. Math. **19**, 65 (1988).
9. D. Bonchev, J. Mol. Struct. (Theochem) **336**, 137 (1995). D. Bonchev and N. Trinajstić, J. Chem. Phys. **67**, 4517 (1977). D. Bonchev and N. Trinajstić, Int. J. Quantum Chem., Quantum Chem. Symp. **12**, 293 (1978). M. Randić, Acta Chim. Sloven. **44**, 57 (1997).
10. M. A. Johnson and G. M. Maggiora, *Concepts and Applications of Molecular Similarity* (Wiley, New York, 1990). M. Randić, in *Mathematical Methods in Contemporary Chemistry*, S. I. Kuchanov, Ed. (Gordon & Breach, New York, 1996). P. G. Mezey, J. Math. Chem. **11**, 27 (1992). S. C. Basak, S. Bertelsen, and G. D. Grunwald, J. Chem. Inf. Comput. Sci. **34**, 270 (1994). M. Randić, J. Chem. Inf. Comput. Sci. **36**, 1092 (1996).
11. M. Randić, J. Chem. Ed. **69**, 713 (1992).
12. Lord Kelvin, *Baltimore Lecture on Molecular Dynamics and the Wave Theory of Light* (C. J. Clay, London, 1904), pp. 439, 619.
13. A. Radecki, H. Lamparczyk, and R. Kaliszan, Chromatographia **12**, 597 (1979).
14. I. Motoc, Z. Naturforsch. A **38**, 1342 (1983). A. J. Hopfinger, J. Am. Chem. Soc. **102**, 7196 (1980).
15. T. R. Stouch and P. C. Jurs, J. Chem. Inf. Comput. Sci. **26**, 4 (1986). R. H. Rohrbaugh and P. C. Jurs, Anal. Chim. Acta **199**, 99 (1987).
16. L. B. Kier, Quantum Struct.-Act. Relat. **4**, 109 (1985). L. B. Kier, Quantum Struct.-Act. Relat. **5**, 1 (1986). L. B. Kier, Quantum Struct.-Act. Relat. **5**, 7 (1986). L. B. Kier, Acta Pharm, Jugosl. **36**, 171 (1986). L. B. Kier, Med. Res. Rev. **7**, 417 (1987).
17. M. Randić, Croat. Chem. Acta **66**, 289 (1993).
18. M. Randić, J. Chem. Inf. Comput. Sci. **35**, 373 (1995).
19. M. Randić, New J. Chem. **19**, 781 (1995). M. Randić, Int. J. Quantum Chem., Quantum Biol. Symp. **22**, 61 (1995). M. Randić, New J. Chem. **20**, 1001 (1996).
20. M. Randić, J. Math. Chem. **19**, 375 (1996).
21. M. Randić and G. Krilov, Int. J. Quantum Chem., Quantum Biol. Symp. **23**, 127 (1996).
22. M. Randić, in *From Chemical Topology to Three-Dimensional Geometry*, A. T. Balaban, Ed. (Plenum Press, New York, 1997).
23. I. I. Baskin, E. V. Gordeeva, R. O. Devdariani, N. S. Zefirov, V. A. Palyulin, and M. I. Stankevich, Dokl. Acad. Nauk SSSR **307**, 613 (1989). M. I. Skvortsova, I. I. Baskin, O. L. Slovokhotova, V. A. Palyulin, and N. S. Zefirov, J. Chem. Inf. Comput. Sci. **33**, 630 (1993).
24. L. B. Kier and L. H. Hall, Quantum Struct.-Act. Relat. **12**, 60 (1993). I. Lukovits, J. Chem. Inf. Comput. Sci. **34**, 287 (1994).
25. V. Kvasnička and J. Pospichal, J. Chem. Inf. Comput. Sci. **36**, 516 (1996).
26. N. S. Zefirov, V. A. Palyulin, and E. V. Radchenko, Doklady Akad. Nauk SSSR **316**, 921 (1991).

Density Functional Calculations of Dioxygen Binding in Mono- and Dinuclear Copper Complexes

ATTILA BÉRCES

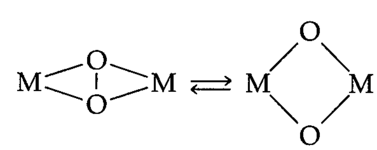
Steacie Institute for Molecular Sciences, National Research Council, Canada, 100 Sussex Dr., Ottawa Ontario K1A 0R6, Canada

Received 3 March 1997; revised 19 June 1997; accepted 23 June 1997

ABSTRACT: We studied the electronic structure, electronic excitation energies, charge distribution, and magnetic properties of three dinuclear copper complexes, $[\{[L\text{-}Cu]_2O_2\}^{2+}$, L = 1,4,7-triaza-cyclo nonone] (**1**), $[\{[L\text{-}Cu]_2O_2\}^{2+}$, L = hydrotis-(pyrazolyl)-borate (**2**) and $\{[(NH_3)_3\text{-}Cu]_2O_2\}^{2+}$ (**3**) in two core isomers. The theoretical model complex **3** is sufficient to describe the qualitative electronic structure and related properties of **1** and **2**. The main features of the electronic structure are similar between the three systems; the electronic excitation energies between copper- and oxygen-based orbitals are insensitive to ligand effects. In addition, we studied the energetics of mononuclear complexes to determine the mechanism of the formation of **1**. In the suggested mechanism the end-on mononuclear complex forms first followed by an isomerization between the trans-μ-1,2-O_2, μ-η^2:ν^2-O_2 and (μ-O)$_2$ dinculear isomers.
© 1997 John Wiley & Sons, Inc. Int J. Quant Chem **65**: 1077–1086, 1997

Introduction

Recently, much attention has focused on metalloenzymes responsible for cleaving and forming the dioxygen O—O bond as well as to those which bind dioxygen reversibly. Most of these enzymes contain one, two, or more metal centers in their active sites with common core structures in their oxygenated form of either $M_2(\mu\text{-}\eta^2{:}\eta^2\text{-}O_2)$, **A**, or $M_2(\mu\text{-}O)_2$, **B** where M = Mn, Fe, or Cu. A tetranculear mangenese cluster which consists of two $Mn_2(\mu\text{-}O)_2$ units (structure **B**) is the active site of photosystem II, which is responsible for the dioxygen evolution from water. The reverse reaction, the cleavage of the dioxygen O—O bond occurs at dinuclear iron and copper sites of several enzymes with a core structure **A**.



A **B**

Based on the structural similarity between **A**

Contract grant sponsor: National Research Council, Canada.

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061077-10

and **B**, it seems reasonable that the general mechanism for oxygen evolution and dioxygen activation in metalloenzymes involves the interconversion between **A** and **B** [1]. Tolman and co-workers recently reported the synthesis of *N*-substituted derivatives of [$\{[L\text{-}Cu]_2O_2\}^{2+}$, L = 1,4,7-triazacyclononane] [2] (**1**) which reversibly bind dioxygen and reversibly break the O—O bond through the isomerization of their $Cu_2(\mu\text{-}\eta^2{:}\eta^2\text{-}O_2)$, **A**, and $Cu_2(\mu\text{-}O)_2$, **B** core structures [3]. One intriguing aspect of this discovery was that other dinuclear copper complexes with core structure **A** have similar structural, magnetic, and spectroscopic properties, but their chemistry is different. Of particular interest are the *N*-substituted derivatives of [$\{[L\text{-}Cu]_2O_2\}^{2+}$, L = hydrotis(pyrazolyl)borate] (**2**) synthesised by Kitajima et al. [4].

In the recent study we showed that the binding energy of dioxygen of **1** and **2** are clearly different, namely −60 and −184 kJ/mol, respectively [5]. These results explained why **1** binds dioxygen reversibly and **2** irreversibly. Further we have shown that the isomerization energetics of **1** and **2** are similar in spite of the different binding energy. Chemical properties are determined by the total energies of the complex to which the ligands make significant contribution while structural, spectral, and magnetic properties are more localized and are reminiscent of the chromophore.

The question remains, therefore, if the same level of theory which gave different total binding energies for **1** and **2** can reproduce the similarities in the properties related to the electronic structure of the Cu_2O_2 chromophore. This question is addressed in the present study. Previous theoretical modeling of **1** and **2** was based on a model complex $\{[(NH_3)_3\text{—}Cu]_2O_2\}^{2+}$ (**3**). To study to what extent the theoretical models can be simplified without sacrificing its capability to reproduce properties, we carried out calculations on **3**. In addition we address the binding energetics and properties of mononuclear dioxygen copper complexes related to **1** to examine the binding mechanism of dioxygen.

For the discussion of spectroscopic properties it is also of interest to make reference to observed spectral properties of oxy-hemocyanin and oxy-tyrosinase active sites. **1** and **2** serve as inorganic mimics of the active sites of these two enzymes in core isomer **A** and their structural, magnetic, and spectral properties are also similar [6, 7]. High-resolution X-ray crystallography indicates that the peroxide is perpendicularly bridging the Cu-Cu axis in a symmetrical fashion, which gives rise to a $Cu_2(\mu\text{-}\eta^2{:}\eta^2\text{-}O_2)$ core structure. Strong antiferromagnetic coupling between the two d^9 copper centers through a superexchange mechanism [8] leads to a diamagnetic, electron paramagnetic resonance (EPR) silent and singlet ground state [9]. The O–O stretching frequency is significantly downshifted to 725–760 cm^{-1} from the dioxygen frequency and is close to a typical peroxide stretching frequency. There are two characteristic absorption bands at 350 nm ($\varepsilon = 20{,}000\ M^{-1}\ cm^{-1}$) and 550 nm ($\varepsilon = 1000\ M^{-1}\ cm^{-1}$) and a feature in the circular dichroism spectrum at 480 nm ($\Delta\varepsilon = +2.5\ M^{-1}\ cm^{-1}$) with no corresponding feature in the absorption spectrum. The Cu–Cu, O–O, and Cu–O distances are 3.6, 1.41, 1.9 Å, respectively.

As one of the first theoretical studies, Solomon and co-workers carried out self-consistent scattered wave X_α (SCF-X_α-SW) calculation [9] to interpret the peroxide → Cu charge-transfer transitions using Slater's transition-state method [10]. Most recently Tuczek and Solomon presented a valence bond configuration interaction (VBCI) model to explain the charge-transfer energy splitting and excited-state antiferromagnetism in bridged transition-metal dimers with applications to oxo-Hc [9c].

Recently, Bernardi and co-workers reported ab initio and density functional calculations on oxy-hemocyanin models including **3** [11] and on the binding process of triplet oxygen to hemocyanin [12]. Eisenstein and co-workers have applied extended Hückel theory (EHT) [13] to a model system with two copper cations ligated by six imidazol rings and bridged by an oxygen molecule (peroxide). This study has been recently extended by Getlicherman et al. to the eclipsed arrangements of the ligands which have been found isoenergic with the staggered conformation [14].

Theoretical studies of the $Cu_2(\mu\text{-}O_2)$ complexes have appeared recently, prompted by the discovery of this core isomer. Mahapatra and co-workers have complemented the experimental work by minimum basis set restricted Hartree–Fock (HF) calculation and by broken symmetry X_α method on **3** [2b]. The broken symmetry calculations converged to the symmetrical solution and the calculated Cu–Cu and O–O distances and O–O vibrational frequency (from HF calculations) were in reasonable agreement with the experimental data of the related inorganic complex. Cramer and co-workers described the core isomerization of **3** by

minimal basis set HF geometry optimization followed by CASPT2 single point calculations [15].

In this study we apply density functional theory (DFT) at generalized gradient approximation level (GGA) to calculate the properties of **1**, **2**, and **3** (see Computational Details). Gradient-corrected DFT has been shown to provide results in excellent agreement with high-level ab initio calculations on simple theoretical model systems of hemocyanin. Further this method includes dynamical electron correlation which was neglected in most previous theoretical studies (HF, CASSCF, X_α, EHT), and which was shown to be essential for the proper description of these systems. Density functional theory has been a remarkably successful tool in modeling energetics, structures, and spectroscopy of transition metal systems [16].

Computational Details

The reported calculations were carried out using the Amsterdam Density Functional (ADF) program system version 2.0.1 derived from the work of Baerends et al. [17] and developed at the Free University of Amsterdam [18] and at the University of Calgary [19]. All optimized geometries in this study were calculated based on the local density approximation [20] (LDA) augmented with gradient corrections to the exchange [18b] and correlation [18c] potentials. The geometries were optimized based on the direct inversion of iterative subspace for geometry (GDIIS) technique [21] using natural internal coordinates [22]. We have combined the ADF program with the GDIIS program [23] and previously implemented the skeletal internal coordinates [24]. The internal coordinates were generated by the INTC program [25] and augmented by hand.

The atomic orbitals on copper were described by an uncontracted triple-ζ Slater-type orbitals (STO) basis set [26], while a double-ζ STO basis set was used for carbon, nitrogen, oxygen, and hydrogen; a single-ζ polarization function was used on all atoms. The $1s^2$ configuration on carbon, nitrogen, and oxygen as well as the $1s^2 2s^2 2p^6$ configuration of copper were assigned to the core and treated by the frozen-core approximation. A set of auxiliary s, p, d, f, g, and h STO functions, centered on all nuclei, was used in order to fit the molecular density and represent the Coulomb and exchange potentials accurately in each SCF cycle [27].

DINUCLEAR COMPLEXES

Core Structure

For the successful prediction of the chromophore properties, it is essential that the calculation reproduces the experimental geometrical parameters. The calculated Cu–Cu distance of isomer **A** of **1** and **2** is 3.70 Å while the experimental counterpart is 3.6(1) Å. The O–O distances of **1A** and **2A** are 1.47 and 1.46 Å, respectively, while the experimental value is 1.41 Å. Considering the challenging task of calculating geometries of bridged dinuclear metal systems and the large experimental error bars, the agreement is good between theory and experiment. More details on the geometry is given in Ref. [5].

ELECTRONIC STRUCTURE

Solomon et al. studied the electronic structure and electron transition energies of side-on and end-on bonded dioxygen complexed to copper monomer and dimer complexes by broken symmetry SCF-X_α-SW calculations [9a]. One important question is whether the electronic structure is fully delocalized or a localized broken symmetry state. The calculations on **3A** by Solomon et al. converged to a broken symmetry solution, but in the series of dicopper peroxo complexes considered, this system had the most delocalized electron spin density indicating the most strongly coupled system. We found that broken symmetry gradient-corrected density functional calculations on the **A** isomer of **1**, **2**, and **3** converged to the symmetrical, fully delocalized solution on all model systems. Our results indicate that the ground state of oxy-hemocyanin mimics can be well described by a restricted single determinant molecular orbital (MO) model. We note here that other functionals may yield broken symmetry solutions especially those which contain mixed Hartree–Fock exchange terms.

The delocalized MOs are significant for the technical details of the calculation, but it does not change the qualitative orbital descriptions fundamentally. In fact, most features of the orbital interactions and orbital energy levels are very similar to that based on broken symmetry SCF-X_α-SW calculations. Since the qualitative orbital interac-

tions have previously been discussed by Solomon et al. [9], and Bernardi et al. [11] on the basis of SCF-X$_\alpha$-SW and DFT calculations, respectively, here we restrict our discussion to the orbital correlation between the two core isomers and to the comparison between the different complexes. Figure 1 shows the correlation diagram between the two core isomers of **1**, **2**, and **3**. These orbital diagrams are shifted so that the highest occupied molecular orbital (HOMO) of isomer **A** is at the same level for all three models. Our symmetry designations refers to C_{2h} symmetry in standard orientation, z axis along the O—O bond (see Fig. 2).

The Π^*_α orbital is strongly stabilized by the bonding interaction with the copper-based d_{xz} orbital which forms the $7B_g$ of **3**. The lowest unoccupied molecular orbital (LUMO) is the antibonding

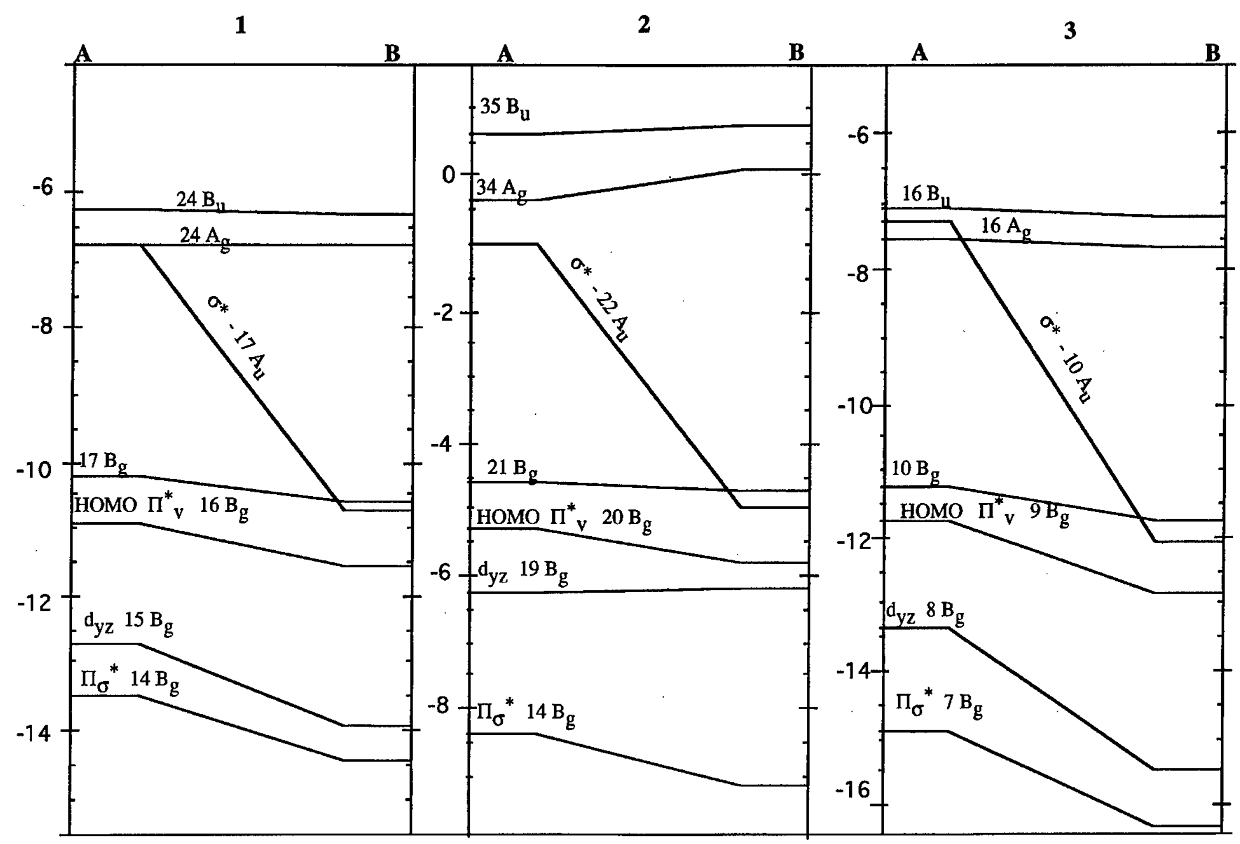


FIGURE 1. Orbital correlation diagram of core isomers.

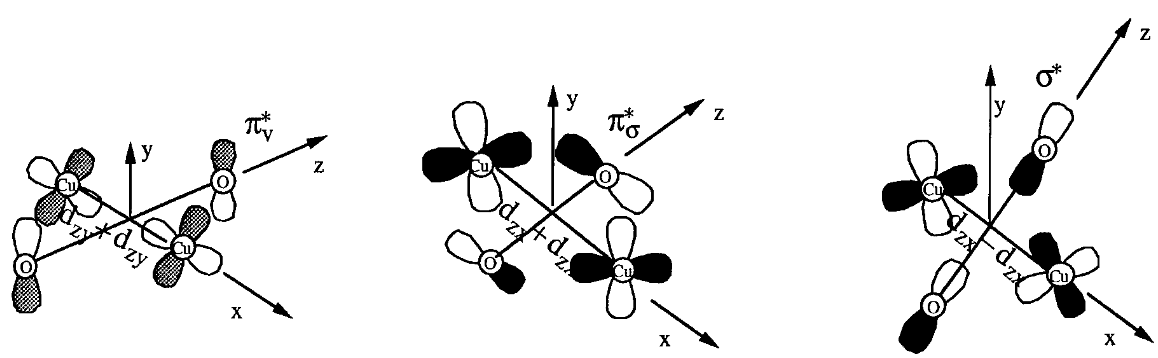


FIGURE 2. Orbital interactions.

combination of the same two orbitals. In our calculation the HOMO ($9B_g$ for **3**) is a primarily Π_v^* orbital with some antibonding mixing of the copper based d_{yz} orbitals, The corresponding bonding orbital ($8B_g$ of **3**) has mainly d_{yz} character. Our description of the HOMO is essentially different from that by the SCF-X_α-SW calculation of Solomon et al [9a]. They found that the d_{xy} orbitals were stabilized by the σ^* of the peroxide which forms the HOMO orbital of the SCF-X_α-SW wave function. The corresponding $9A_u$ orbital in **3** (not shown) is almost 1 eV below the HOMO of our calculation. Further, our calculation did not show any indication that the peroxide σ^* acts as a π acceptor orbital for core isomer **A** in contrast to such finding by SCF-X_α-SW calculations. However, such interaction is an essential feature of core isomer **B**. In fact the most striking difference between the electronic structures of **A** and **B** isomers is the strong stabilization of the σ^* orbital in isomer **B**. All MOs containing Π^* orbitals of the peroxide also drop in energy, while the $16B_u$ and the $16A_g$ of **3** consisting of the appropriate combinations of Cu based s orbitals do not change in energy.

Comparing the MO diagram of the different systems, **3** is a good qualitative representation of the electronic structure of the inorganic mimics. We compared the linear combination coefficients of **1**, **2**, and **3**, and found remarkable similarities in most oxygen- and copper-based orbitals. Most of the differences are related to the extent ligand orbitals participate in the binding.

Electronic Excitation Energies

We also compared the electronic excitation energies between the three model systems. Solomon et al. [9] have shown that configuration interactions, which are not included in our calculations, can contribute as much as 20,000 cm^{-1} to the excitation energies of systems with Cu_2O_2 core. Therefore, our goal is to test the theoretical system, **3**, of its ability to reproduce single-electron excitation energies of the inorganic systems **1** and **2**. We calculated the two optically allowed $\pi^* \rightarrow$ Cu(II) ligand to metal charge-transfer transitions of isomer **A** of **1** and **3** based on the sum method developed by Ziegler et al. [28].

The calculated π_v^* transition for **1** and **3** are 36,000 and 37,500 cm^{-1}, respectively. Similarly, the π_σ^* transition energies are also relatively close in **1** and **3**; these are 65,200 and 61,200 cm^{-1}, respectively. However, both the π_v^* and π_σ^* transition energies are about twice the experimental values of 18,000 and 28,600 cm^{-1}, respectively. SCF-X_α-SW transition energies obtained by Slater's transition state method are 16,000 cm^{-1} and 66,800 cm^{-1} for the π_v^* and π_σ^* transitions, respectively.

Magnetic Properties

The ground-state Heisenberg magnetic coupling constant ($-2J$) for a strongly coupled dinuclear magnetic system is simply the difference between the ground singlet and triplet state energies [9]. Experimentally, hemocyanin and the inorganic mimics studied here are all diamagnetic. Only the lower limit of the $-2J$ constant has been determined experimentally, which is 600 cm^{-1}. SCF-X_α-SW calculations predict the value of the antiferromagnetic coupling between 5675 and 11800 cm^{-1} [9]. The gradient-corrected DFT calculations on **3** by Bernardi et al. [11] yields 13.34 kcal/mol or 4665 cm^{-1}. Our calculated $-2J$ values of **1** and **3** are 6163 and 3800 cm^{-1}, respectively, while the $-2J$ value of core isomer **B** is slightly lower, 3387 cm^{-1} for **1**, for example.

Charge Analysis

We also investigated the charge transfer from the Cu(I)L to the dioxygen in the complex. Table I contains the calculated partial charges on Cu and O calculated by three different charge analysis. The Mulliken population analysis is the most

TABLE I
Partial charges on Cu and O.

μ-η^2:η^2-O_2	Method	1	2	3
Cu	Mulliken	0.77	0.72	0.71
Cu	Hirshfeld	0.42	0.40	0.42
Cu	Vornoi	0.88	0.95	0.91
O	Mulliken	−0.50	−0.29	−0.45
O	Hirshfeld	−0.22	−0.20	−0.19
O	Vornoi	−0.46	−0.45	−0.41
μ-O_2	Method	1	2	3
Cu	Mulliken	0.91	0.94	0.93
Cu	Hirshfeld	0.46	0.46	0.49
Cu	Vornoi	1.03	1.07	1.08
O	Mulliken	−0.79	−0.77	−0.76
O	Hirshfeld	−0.35	−0.31	−0.30
O	Vornoi	−0.70	−0.68	−0.63

straightforward, but its physical meaning is somewhat questionable. The Hirshfield analysis provides partial atomic charges as an integral of the SCF charge density over space, in each point weighted by the relative fraction of the initial density of that atom in the total initial density [29]:

$$q^{\mathrm{atom}(i)} = \int \frac{\rho_{\mathrm{atom}(i)}^{\mathrm{initial}}}{\sum_j \rho_{\mathrm{atom}(j)}^{\mathrm{initial}}} \rho^{\mathrm{SCF}}. \tag{1}$$

The Voronoi charge analysis consists of assigning the charge density in a point in space to the nearest atom. The Voronoi cell of an atom is the region in space closer to that atom than any other.

In spite of the different ligand charges, the calculated partial charges on the Cu and O atoms of the complexes are very similar in all systems. This is in agreement with what is expected based on the similarity in the core electronic structure. The Hirshfeld analysis also allows one to directly calculate the amount of charge transferred from one fragment to the other by applying Eq. (1) to fragments rather than atoms. On average the total charge transferred from two mononuclear Cu(I)L complexes to the dioxygen is -0.65 and -1.02 for the μ-η^2:η^2-O_2 and the μ-O_2 isomer, respectively, which indicates a high degree of covalency in these systems.

MONONUCLEAR SUPEROXO COMPLEXES AND THE MECHANISM OF OXYGEN BINDING

It is difficult to study experimentally or theoretically the binding mechanism of dioxygen in the biological system. However, there is adequate experimental data about the bonding mechanisms in inorganic complexes, and therefore it is of interest to study the binding mechanisms in these systems. Especially interesting are the systems with triazacyclononame (TACN) ligands. To gain information about the mononuclear intermediates of the binding process, we optimized the geometry of both the side-on (**4a**) and end-on (**4b**) isomers of the mononuclear superoxo adducts of Cu(I)L^+ systems with TACN ligands. Experimentally the side-on bonded superoxo adducts were observed with tridentate hydrotris (3-*t*-butyl-5-isopropyl-pyrazolyl)borate ligand [30] and terminal end-on complexes with the tetradentate tris(2-pyrydylmethyl)amine and related ligands [31]. Based on these observations and other related studies on expects that the tridentate TACN ligand also gives a stable side on bonded superoxo adduct. Experimentally, however, such complex has not been isolated probably because the dinuclear complex is the thermodynamic product. Theoretical calculations make it possible to characterize these transient species and the energetic comparisons provides insight into the binding mechanism.

The calculated energetic and structural information of the optimized structures is listed in Table II and the optimized geometries are shown in Figure 3. The geometry of the side-on bonded monomer is similar to that of the appropriate part of the dinuclear compound. There are major differences in the O—O and Cu—O bond lengths; the O—O bond is much shorter, and the Cu—O bond is longer than in the dinuclear compound. The apical Cu—N bond length is between that of the Cu(I)L^+ monomer and the dinuclear compound. The end-on superoxo complex has the shortest O—O bond of 1.28 Å and the Cu—O bond length is close to that in the μ-O_2 dinuclear complexes. The conformation of the TACN ligand in the end-on isomer is different from that of the dinuclear adduct and the side-on isomer. There is one short and two long Cu—N bonds similarly to the monomeric Cu(I)L system. However, the (N-)H atoms are almost eclipsed with one of the hydrogens on the neighboring carbon atoms in the side-on conformer. This conformation would be less likely with bulky *N*-sub-

TABLE II
Geometrical and energetic parameters of mononuclear O_2 adducts.[a]

	C_s side-on	C_1 side-on	C_1 end-on
O_2 binding energy[b]	−89	−94	−81
R_{OO}	1.310	1.311	1.281
R_{CuO1}	2.071	2.087	1.896
R_{CuO2}	2.071	2.047	2.786
R_{CuN1}	2.170	2.250	2.064
R_{CuN2}	2.088	2.070	2.193
R_{CuN3}	2.088	2.055	2.135
α	82.7	88.0	81.6
β	118.0	113.3	
γ	22.5	23.3	
O-Cu-Cu-N_1 dihedral	90.0	95.6	

[a] Distances in angstrom, angles in degrees. Numbering of N atoms: 1 is apical, 2 and 3 are equatorial atoms. α, β, γ are N_i-Cu-N_j bending angles with i, j = (2, 3); (1, 2) and (1, 3), respectively.
[b] Reaction energy for O_2 + Cu(I)L → O_2CuL.

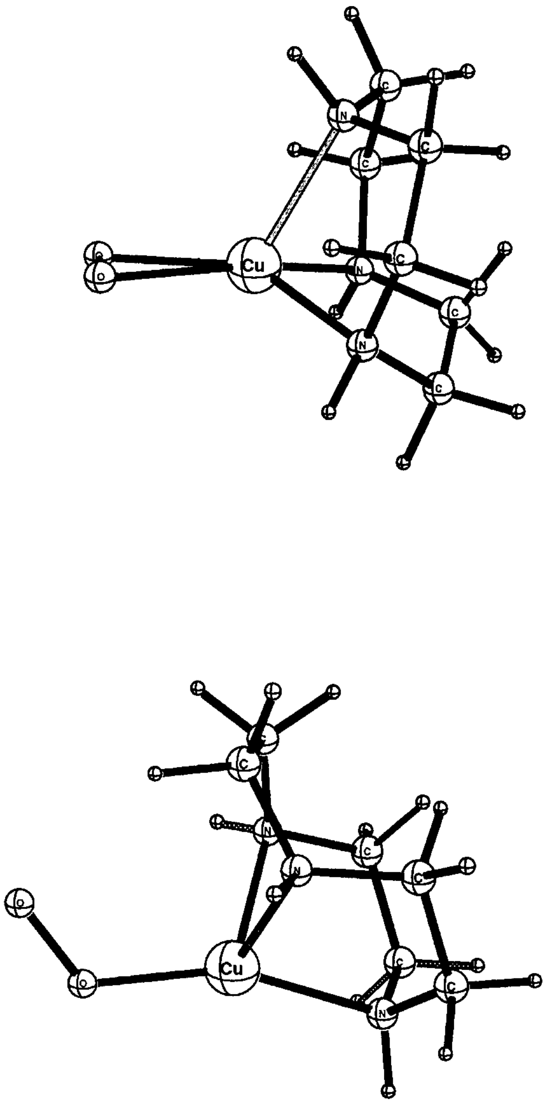


FIGURE 3. Optimized geometries of superoxo complexes.

stituents. Therefore, our calculations most likely overestimate the stability of the side-on conformer compared to the inorganic systems they represent.

The binding energies of dioxygen to Cu(I)L monomers are -94 and -81 kJ/mol for **4a** and **4b**, respectively. In comparison, the binding energy of the dinuclear complex **1** is -60 kJ/mol. The five coordinate side-on adduct **(4a)** is expected to be lower in energy than the four coordinate end-on **(4b)** structure due to the tendency of Cu(II) for five coordination. However, the higher stability of the mononuclear vs. dinuclear system is in contrast to the experimental results, where the dinuclear complex was the thermodynamically stable product. This disagreement is most likely due to simplifications in our computational models and assumptions rather than error in the theory. Most importantly, the *N*-substituents may reverse the stability of the mono- and dinuclear systems. The reversal of stability upon ligand substitution between mononuclear vs. dinuclear copper dioxygen complexes have been experimentally demonstrated. Karlin et al. [32] studied the kinetics and thermodynamics of the oxygenation of copper(I) complexes with tripodal tetradentate ligands: tris[(2-pyridyl)methyl]amine (TMPA) and corresponding ligands with one (BPQA), two (BQPA), or three (TMQA) 2-quinolyl groups substituting for the 2-pyridyl donors. The dioxygen binding enthalpies of the dinuclear complexes were between -50 and -80 kJ/mol, and the relative thermodynamic stability of the mononuclear and dinuclear complexes are different with different ligands. With the BPQA ligand the mononuclear adduct is not observed, while the BQPA system gives the mononuclear complex as the thermodynamically stable product and the dinuclear complex as an intermediate. The binding enthalpy for the mononuclear adduct was measured to be -34 kJ/mol.

Nonetheless, the characterization of an end-on coordinated mononuclear copper complex with tridentate ligand is an important outcome of this study. Experimentally only end-on bonded superoxo complexes with the tetradentate TMPA-analog ligands with 6-pivalolylamine substituents on the pyridyl groups have been structurally characterized [31]. For the side-on adduct our calculations predict strong dioxygen binding, but this species may not be observed due to large kinetic barrier at low temperature and unfavorable entropies at high temperature. Considering these findings and what is experimentally known about these systems, we can comment on the mechanism of dioxygen binding of $CuTACN^+$ and related systems. Experimental results suggest a mechanism which involves the rate-determining formation of a mononuclear superoxo adduct followed by the trapping of the second monomeric Cu complex and swift equilibrium between the two core isomers. The mononuclear adduct cannot be observed as a stable intermediate and with access of O_2 the kinetics is first order in Cu concentration.

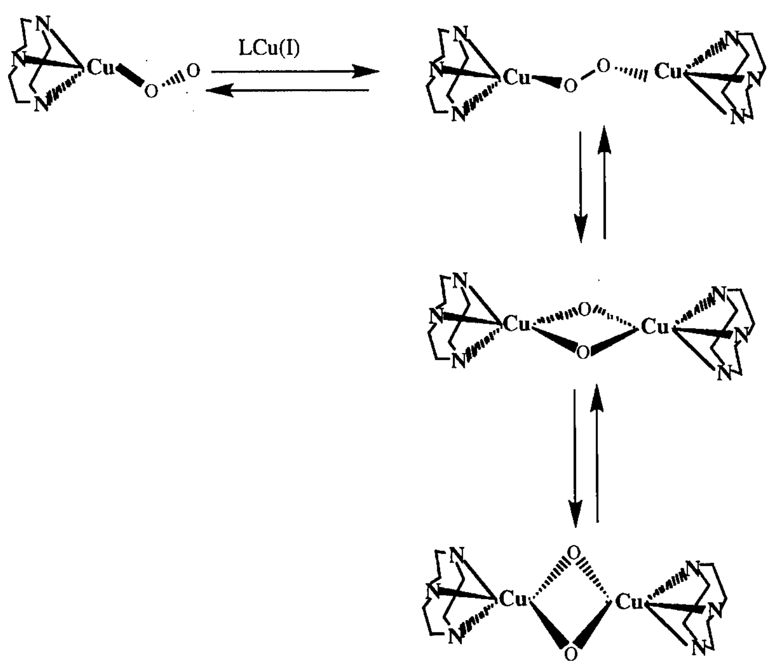
Another structural finding which helps in the interpretation of the binding mechanism is the distortion of the core into an asymmetrical position in the C_i symmetry conformers of the dinuclear

TACN complex. Further distortion along the same coordinate would result in a trans-μ-1,2-O_2 dinuclear complex. This finding suggests that the μ-η^2:η^2-O_2 and trans-μ-1,2-O_2 isomers may also convert to one another. Experimentally the trans-μ-1,2-O_2 dicopper complexes are well known [33]. Further, the formation of a trans-μ-1,2-O_2 dinuclear system from a side-on complex is more sterically favored compared to the end-on monomer. Therefore, we suggest that the mechanism of O_2 binding occurs through the formation of a side-on mononuclear adduct with triplet ground state, followed by the formation of a trans-μ-1,2-O_2 dinuclear complex which rearranges into a μ-η^2:η^2-O_2 complex. The trans-μ-1,2-O_2 dinuclear complex may be the transition state or a short-lived transient form. This mechanism is shown on Scheme 1.

The binding mechanism of the inorganic mimic may be significantly different from that of the biological system. However, one key element of this mechanism is the interconversion between the trans-μ-1,2-O_2 and the μ-η^2:η^2-O_2 core isomers, which was previously suggested by Ling et al. [34] on the basis of resonance Raman spectroscopy of oxy-Hcs and corresponding normal coordinate analysis. Further, such mechanism is consistent with the geometric data about the deoxy-Hc and trans-μ-1,2-O_2 dinuclear complexes. The typical Cu–Cu distance in end-on trans μ-1,2 copper–peroxide complex $[(Cu(I)L)_2(O_2)]2^+$ is 4.36 Å [33], which is somewhat shorter than 4.6 Å, which is found experimentally in one form of deoxyhemocyanin from *Limulus polyphemus* [35]. The constraints set by the protein environment clearly have an effect on the final mechanism of O_2 binding. Cruse et al. [36] have shown that forcing the two copper atoms into close proximity by the presence of a bridging ligand causes a dramatic enhancement of the rate of reaction with O_2, while the reaction enthalpy does not change significantly compared to the nonbridged system.

Conclusions

We compared the electronic structure, electronic excitation energies, charge distribution, and magnetic properties of two dinuclear copper complexes and the corresponding theoretical model complex with ammonia ligands. We found that all properties which are characteristic to the Cu_2O_2 chromophore are similar in all three systems. Since all molecular orbitals involved in the electronic transitions are either copper or oxygen based, the electronic excitation energies are similar for the three system. The $(-2J)$ magnetic constant is the most sensitive property to ligand effects due to ligand participation in the LUMO, but ligand substitution



SCHEME 1. Mechanism of dioxygen binding and O—O bond cleavage.

does not change the diamagnetic property of the complex.

To determine the mechanism of dioxygen binding we studied the mononuclear complexes with triazacyclononane ligand. The mononuclear complexes were found more stable than the dinuclear one, which is most likely related to the simplification in the model, especially the effect of *N*-substitution could reverse the stability as indicated by experiment. We suggest that side-on mononuclear complex formation is followed by binding another monomer in a trans-μ-1,2-O_2 complex. The trans-μ-1,2-O_2 then transforms into μ-η^2:η^2-O_2 core isomers which is an equilibrium with the (μ-O)$_2$ isomer.

ACKNOWLEDGMENTS

I thank Professor Peter Pulay for providing the INTC program, which was used to generate the input for natural coordinate optimization. I gratefully acknowledge a visiting fellowship from the National Research Council Canada. I also thank my colleagues at the Steacie Institute for discussions; I am especially grateful to Dr. Heinz-Bernhard Kraatz for discussions on the chemistry of these systems and to Dr. Allen East for his comments on the manuscript. I thank Shelly Pinder (University of Lethbridge) for the initial calculations. This paper is issued as NRCC# 40820.

References

1. (a) D. M. Proserpio, R. Hoffman, and G. C. Dismukes, J. Am. Chem. Soc. **114**, 4374 (1992). (b) A. L. Feig and S. J. Lippard, Chem. Rev. **94**, 759 (1994). (c) V. L. Pecoraro, M. J. Baldwin, and A. Gelasco, Chem. Rev. **94**, 807 (1994).
2. (a) S. Mahapatra, J. A. Halfen, E. C. Wilkinson, G. Pan, C. J. Cramer, L. Que Jr., and W. B. Tolman, J. Am. Chem. Soc. **116**, 9785 (1995). (b) S. Mahapatra, J. A. Halfen, E. C. Wilkinson, G. Pan, C. J. Cramer, L. Que, Jr., and W. B. Tolman, J. Am. Chem. Soc. **117**, 8865 (1995).
3. J. A. Halfen, S. Mahapatra, E. C. Wilkinson, S. Kaderli, V. G. Jr. Young, L. Que Jr., A. D. Zuberbühler, and W. B. Tolman, Science **271**, 1397 (1996).
4. (a) N. Kitajima, T. Koda, S. Hashimoto, T. Kitagawa, and Y. Moro-oka, J. Chem. Soc. Chem. Commun. 151 (1988). (b) N. Kitajima, K. Fujisawa, Y. Moro-oka, and K. Toriumi, J. Am. Chem. Soc. **111**, 8975–8976 (1989).
5. A. Berces, submitted.
6. (a) K. Magnus, H. Ton-That, and J. E. Carpenter, in *Bioinorganic Chemistry of Copper*, K. D. Karlin, Z., Tyeklar, Eds. (Chapman and Hall, New York, 1993), pp. 143–150. (b) K. Magnus, H. Ton-That, and J. E. Carpenter, Protein: Struct. Funct. Genet. **19**, 302 (1994).
7. A. Volbeda and G. J. Hol, J. Mol. Biol. **209**, 249 (1989).
8. (a) P. W. Anderson, Solid State Phys. **14**, 99 (1963). (b) P. J. Hay, J. C. Thiebault, and R. Hoffman, J. Am Chem. Soc. **97**, 4884 (1975). (c) K. Yamaguchi, Y. Takahara, T. Fueno, and K. Nasu, Jpn. J. Appl. Phys. **26**, L2307 (1987). (d) P. DeLoth, P. Cassoux, and J.-P. Daudey, J. Am. Chem. Soc. **103**, 4007 (1981). (e) M. Gerloch, in *Progress in Inorganic Chemistry*, S. Lippard, Ed. (Interscience, New York 1979).
9. (a) P. K. Ross and E. I. Solomon, J. Am Chem. Soc. **112**, 5871 (1990). (b) P. K. Ross and E. I. Solomon, J. Am. Chem. Soc. **113**, 3246 (1991). (c) F. Tuczek and E. I. Solomon, J. Am. Chem. Soc. **116**, 6916 (1994).
10. J. C. Slater, *The Self-consistent Field for Molecules and Solids* (McGraw-Hill, New York, 1974), Vol. 4.
11. F. Bernardi, A. Bottoni, R. Casadio, P. Fariselli, and A. Rigo, Int. J. Quant. Chem. **58**, 109 (1996).
12. F. Bernardi, A. Bottoni, R. Casadio, P. Fariselli, and A. Rigo, Inorg. Chem. **35**, 5207 (1996).
13. O. Eisenstein, C. Giessner-Prettre, J. Maddaluno, D. Stussi, and J. Weber, Arc. Biochem. Biophys. **296**, 247 (1992).
14. H. Getlicherman, C. Giessner-Prettre, and J. Maddaluno, J. Phys. Chem. **100**, 6819 (1996).
15. (a) S. Mahapatra, J. A. Halfen, E. C. Wilkinson, G. Pan, X. Wang, V. G. Young Jr., C. J. Cramer, L. Que, Jr, and W. B. Tolman, J. Am. Chem. Soc. **118**, 11555 (1996). (b) C. J. Cramer, B. A. Smith, and W. B. Tolman, J. Am. Chem. Soc. **118**, 11283 (1996). (c) B. A. Smith and C. J. Cramer, J. Am. Chem. Soc. **118**, 5490 (1996).
16. T. Ziegler, Chem. Rev. 91, 651 (1991).
17. E. J. Baerends, D. E. Ellis, and P. Ros, Chem. Phys. **2**, 41 (1973).
18. (a) W. Ravenek, in *Algorithms and Applications on Vector and Parallel Computers*, H. J. J. te Riele, Th. J. Dekker, H. A. van de Vorst, Eds. (Elsevier, Amsterdam, 1987). (b) P. M. Boerrigter, G. te Velde, and E. J. Baerends, Int. J. Quant. Chem. **33**, 87 (1988). (c) G. te Velde, and E. J. Baerends, J. Comp Phys. **99**, 84 (1992).
19. L. Fan and T. Ziegler, J. Chem. Phys. **94**, 6057 (1991); L. Fan and T. Ziegler, J. Chem. Phys. **95**, 7401 (1991). (a) L. Fan, L. Versluis, T. Ziegler, E. J. Baerends, and W. Ravenek, Int. J. Quant. Chem. **S22**, 173 (1988). (b) L. Fan and T. Ziegler, J. Chem. Phys. **96**, 9005 (1992). (c) L. Fan and T. Ziegler., J. Phys. Chem. **96**, 6937 (1992).
20. (a) S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys, **58**, 1200 (1980). (b) A. D. Becke, Phys. Rev A. **38**, 2398 (1988). (c) J. P. Perdew, Phys. Rev. B **33**, 8822 (1986); *ibid.*; **B34**, 7046 (1986).
21. (a) P. Csáazár and P. Pulay, J. Mol. Struct. **114**, 31 (1984). (b) P. Pulay, Chem. Phys. Lett. **73**, 393 (1980). (c) P. Pulay, J. Comput. Chem. **3**, 556 (1982).
22. (a) G. Fogarasi, X. Zhou, P. W. Taylor, and P. Pulay, J. Am. Chem. Soc. **114**, 8191 (1992). (b) P. Pulay, G. Foarasi, F. Pang, and J. E. Boggs, J. Am. Chem. Soc. **101**, 2550 (1979).
23. Programmed by A. G. Csaszar and P. G. Szalay, Eötvös University, Budapest, Hungary (1984).
24. (a) A. Bérces, J. Comp. Chem. **18**, 45 (1996). (b) A. Bérces and T. Ziegler, J. Phys. Chem. **98**, 13233 (1994). (c) A. Bérces, T. Ziegler, and L. Fan, J. Phys. Chem. **98**, 1584 (1994).

25. INTC program to generate natural internal coordinates, P. Pulay and G. Fogarasi, 1992, University of Arkansas, Fayetteville.
26. (a) G. J. Snijders, E. J. Baerends, and P. Vernooijs, At. Nucl. Data. Tables **26**, 483 (1982). (b) P. Vernooijs, G. J. Snijders, and E. J. Baerends, Slater Type Basis Functions for the whole Periodic System, Internal report, Free University of Amsterdam, The Netherlands, 1981.
27. J. Krijn and E. J. Baerends, Fit Functions in the HFS-Method, Internal report (in Dutch), Free University of Amsterdam, The Netherlands, 1984.
28. T. Ziegler, A. Rauk, and E. J. Baerends, Theor. Chim. Acta **43**, 261 (1977).
29. (a) F. L. Hirshfield, Theoret. Chim. Acta **44**, 129 (1977). (b) K. B. Wiberg and P. R. Rablen, J. Comp. Chem. **14**, 1504 (1993).
30. K. Fujusawa, M. Tanaka, Y. Moro-oka, and N. Kitajima, J. Am. Chem. Soc. **116**, 12079 (1994).
31. M. Harata, K. Jitsukawa, H. Masuda, and H. Einaga, J. Am. Chem. Soc. **116**, 10817 (1994).
32. K. D. Karlin, N. Wei, B. Jung, S. Kaderli, P. Niklaus, and A. Zuberbüler, J. Am. Chem. Soc. **115**, 9506 (1993).
33. R. R. Jacobson, Z. Tyeklar, A. Farooq, K. D. Karlin, S. Liu, and J. Zubieta, J. Am. Chem. Soc. **110**, 3690 (1988).
34. J. Ling, L. P. Nestor, R. S. Czernuszewicz, T. G. Spiro, R. Frackiewicz, K. D. Sharma, T. M. Loehr, and J. Sanders-Loehr, J. Am. Chem. Soc. **116**, 7682 (1994).
35. B. Hazes, K. A. Magnus, C. Bonaventura, J. Bonaventura, Z. Dauter, K. Kalk, and W. G. Hol, J. Protein **2**, 597 (1993).
36. R. W. Cruse, S. Kaderli, K. D. Karlin, and A. D. Zuberbüler, J. Am. Chem. Soc. **110**, 6882 (1988).

Local Dielectric Constants and Poisson–Boltzmann Calculations of DNA Counterion Distributions

GENE LAMM, GEORGE R. PACK

UIC College of Medicine, 1601 Parkview Ave., Rockford, Illinois 61107

Received 5 March 1997; accepted 6 August 1997

ABSTRACT: Recent work by the authors on the calculation of local solvent dielectric constants around polyelectrolytes using the Poisson–Boltzmann approach is analyzed in terms of the effect on surface potentials and counterion concentrations. Polyelectrolyte surface geometry, local electric fields, and counterion distributions contribute to the self-consistent prediction of local solvent dielectric constants. For an all-atom cell model of DNA with added monovalent salt varying from 0 to 0.5*M*, the Poisson–Boltzmann-determined electrostatic potential increases (negatively) by 50–100% upon the inclusion of local dielectric constants. This, in turn, implies that hydronium ion concentrations in the major and minor grooves increase by about 0.65 and 0.35 pH units, respectively. While counterion concentrations in the major groove change only slightly, those in the minor groove increase by 60–90%. It is also noted that while the local dielectric constant in the major groove monotonically increases away from the surface toward the bulk value of water the dielectric constant in the minor groove has a minimum about 2 Å from the surface due primarily to the local electric field. Certain other properties, such as ionic and dipole first passage times, are affected little by local dielectric constants (less than about 3%). © 1997 John Wiley & Sons, Inc. Int J Quant Chem **65**: 1087–1093, 1997

Correspondence to: G. R. Pack.
Contract grant sponsor: National Institutes of Health.
Contract grant number: GM29079.

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061087-07

Introduction

Recent calculations of thermodynamic properties of polyelectrolytes in solution rely largely on the application of the Poisson–Boltzmann method [1]. The Poisson–Boltzmann (PB) approach belongs to that class of numerical techniques that are both realistic and "computationally convenient." It is realistic in that it can be applied to a model of a polyelectrolyte such as DNA in which hundreds and possibly thousands of atoms are fixed in a reasonable "equilibrium" configuration and the resulting distribution of solvated counterions determined. It is "computationally convenient" in the sense [2] that calculations on systems such as those above can be performed quickly and easily on everyday personal computers.

However, despite these advantages over other popular equilibrium techniques such as the Monte Carlo method, several approximations inherent in most applications of the PB theory limit its predictive capability. Most often mentioned is the traditional neglect of ion correlation and also the cavalier treatments of both the solvent and the solvated counterions in terms a uniform structureless dielectric constant and vanishing ionic radii. The effect of each approximation is difficult to estimate but comparison with well-defined Monte Carlo calculations suggests that the lack of ion correlation yields PB values for monovalent counterion concentrations at the surface of all-atom models of DNA that are about 15–20% low [3–5]. Furthermore, corrections to standard PB theory by describing counterions with nonzero radii [6] but still neglecting ion correlation leads to poorer agreement with Monte Carlo data [7]. Inclusion of correlation within the PB framework is possible [8], but its application to typical all-atom polyelectrolyte models is far less "convenient" than are standard approaches.

The focus of considerable recent work has been on removing the constraint of a uniform solvent dielectric constant. Most modern PB algorithms can be traced back to the work of Warwicker [9–11], which can be easily modified to incorporate a variable local (i.e., solvent) dielectric "field" [12–15]. These latter works have relied on educated guesses about local dielectric values, so estimates of errors involved in specific solvent treatments are qualitative at best. To amend this situation, the present authors presented a method by which local solvent dielectric constants can be calculated within the PB approximation [16]. Comparison with available experiment data shows this self-consistent PB extension to be as accurate as the PB method itself, i.e., it yields local dielectric constants that are about 15–20% too large at the surface of an all-atom representation of DNA. The purpose of the present article is to analyze the effect of including this nonuniform dielectric constant treatment upon the resulting PB local potentials and counterion concentrations. The following section gives a very brief overview of the algorithm. Numerical results and conclusions drawn from them are presented in the final section.

The Calculation of Local Dielectric Constants

The extension of the original Debye–Hückel method [17] (sometimes called the linear PB approximation) to ionic distributions around polyelectrolytes involves describing the electrostatic potential $\phi(\mathbf{r})$ at points in the polyelectrolyte environment by the Poisson equation

$$\nabla \cdot [\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -4\pi\rho(\mathbf{r}), \tag{1}$$

where $\epsilon(\mathbf{r})$ denotes the spatially varying dielectric constant and $\rho(\mathbf{r})$ denotes the charge distribution of both the polyelectrolyte and its ionic environment. The charge distribution may be expressed as a sum of individual ionic species distributions:

$$\rho(\mathbf{r}) = \sum_k \rho_k(\mathbf{r}), \tag{2}$$

where k signifies a particular ionic species. The (normalized) distribution of each species is determined from the electrostatic potential through Boltzmann's equation:

$$\rho_k(\mathbf{r}) = N_k \exp[-\beta z_k \phi(\mathbf{r})] \bigg/ \int d\mathbf{r} \exp[-\beta z_k \phi(\mathbf{r})], \tag{3}$$

where species k has charge z_k and number N_k with $\beta = 1/k_B T$. The solution to coupled Eqs. (1)–(3) for all-atom models of a polyelectrolyte such as DNA is most often accomplished by operating on a three-dimensional grid. The finite-element representation of Eq. (1) on a non-

Cartesian grid is given by

$$\phi_i = \left[4\pi v_i \rho_i + \sum_j (\phi_j \epsilon_{ij} S_{ij}/d_{ij})\right] \Big/ \sum_j (\epsilon_{ij} S_{ij}/d_{ij}), \tag{4}$$

where spatial locations are now indicated by volume element index i and the sum over element j is only over those elements bordering element i. In Eq. (4), v_i is the volume of element i; $\epsilon_{ij} = (\epsilon_i + \epsilon_j)/2$, the arithmetic average of the local dielectric constants of elements i and j; S_{ij}, their shared surface area; and d_{ij}, the distance between their centers. (Simple Cartesian grids have S_{ij} and d_{ij} constant.) The self-consistent iterative solution to Eqs. (2)–(4) for the non-Cartesian grid used by the authors with local dielectric constants fixed at predetermined values (e.g., $\epsilon = 4$ for volume elements representing the DNA and $\epsilon = 78.5$ for bulk water *with* counterions) was described elsewhere [15].

As mentioned above, previous solutions to the finite-element PB equations have limited the description of local dielectric constants to fixed predetermined values [12–15]. It is, however, well known that the relatively high bulk value of the dielectric constant of water is due to the tetrahedral nature of nearest-neighbor bonding [18, 19], which would be interrupted at the surface of a macromolecule. Furthermore, the presence of high electric fields [20] and large counterion concentrations [21] at the surface of polyelectroytes would also be expected to affect local solvent dielectric constants. While surface interruption of tetrahedral bonding is most easily regarded as a "boundary condition" independent of environmental parameters such as ion type or concentration, effects due to electric-field strength and counterion distribution need to be determined self-consistently within the PB iterative algorithm. (Ideally, the effect of ionic concentration would also be taken into account in determining the equilibrium structure of polyelectrolytes such as DNA.) The specific inclusion of these surface, electric field, and ionic distribution effects into the PB approach with a view toward comparing predicted local solvent dielectric constants with experimental data was the subject of a previous article [16]. Here, we are interested in how these effects alter PB-determined potentials and counterion distributions near the surface of the DNA.

To calculate the result of bond disruption at the surface of a macromolecule, we begin with the original Kirkwood result for the bulk dielectric constant of water [18]:

$$\epsilon = 2\pi\beta N_A g \mu^2 + n^2/2, \tag{5}$$

where the g-factor accounting for dipole correlation is given by

$$g = 1 + z\langle \cos\theta \rangle, \tag{6}$$

with z being the number of nearest neighbors of a central molecule, $\langle \cos\theta \rangle$ denoting the average over nearest neighbors of the angle between the molecular dipole vectors, and μ representing the external moment of a spherical sample with macroscopic index of refraction n. For water at 298 K, Kirkwood used $g = 2.64$ and $\mu = 2.15$ D, giving a bulk dielectric value of $\epsilon = 63$. Haggis et al. [22, 23] extended the Kirkwood calculation by explicitly including the tetrahedral bonding of water. The application and modification of their theory to account for tetrahedral bond disruption at surfaces is presented in [16]. The result is that the correlation and total moment of a spherical cluster can be expressed as a sum over cluster sizes:

$$g\mu^2 = \sum_m n_m g_m(f)[\mu_m(f)]^2, \tag{7}$$

where the central molecule of a cluster is bonded to m other water molecules, with m ranging from 0 to 4. In Eq. (7), n_m is the relative population of clusters with bond number m and g_m and μ_m are determined by linearly interpolating between an isolated water molecule ($g_0 = 1$, $\mu_0 = 1.85$ D) and an ice cluster ($g_4 = 2.91$, $\mu_4 = 2.45$ D). Surface effects are represented by the variable f, which denotes the fraction of the volume of a sphere (or cluster) of radius R accessible to the solvent. For a sphere at a distance w from a plane boundary, this fraction is given by

$$f(w) = 0.5 + 0.75w/R - 0.25(w/R)^3. \tag{8}$$

Note that in [16] the phrase prior to the description of Eq. (12), which defines $f(w)$, should read "... we equate the probability that one or more cluster branches is *not* shortened to that fraction of the volume of a sphere of radius R that is accessible to the cluster." For more realistic all-atom models of macromolecules, the fraction f_i at volume elements near the surface (within about 6 Å) needs to be determined numerically. At 298 K, Eqs. (5) and (7) may be combined and simplified to give the

local dielectric constant at element i (in the presence of a boundary) [16]:

$$\epsilon_i^B = 17.5(1 + 1.7f_i^{1/6} + 0.5f_i^{1/3} + f_i + 0.24f_i^{1/2}) + 0.8. \quad (9)$$

The second effect that we want to account for is the diminution of the dielectric constant of water due to dielectric saturation at high electric-field strengths. Booth [20] derived an expression that reproduces the bulk value in the absence of a field. Her result may be combined with Eq. (9) and written [16]

$$\epsilon_i^{BE} = 1.8 + (\epsilon_i^B - 1.8)L(0.08E_i), \quad (10)$$

where ϵ_i^B is the local dielectric constant due to boundary effects only; $L(x)$, the Langevin-type function $3[\coth(x) - 1/x]/x$; and E_i, the electric-field strength in mV/Å at element i (at 298 K). Electric-field strengths are easily determined from potentials using the relation

$$E_i^2 = \sum_j [(\phi_i - \phi_j)/d_{ij}]^2, \quad (11)$$

where the sum is only over neighboring volume elements.

The third effect to be considered is the dielectric decrement of water in the presence of ions. This type of dielectric saturation [21] results from the orientational freezing of water molecules in the solvation shell of an ion and depends on the size of the ion as well as on the number of water molecules the ion "freezes." For monovalent counterions represented by sodium, the local solvent dielectric constant is given by [16]

$$\epsilon_i = \epsilon_i^{BE}[(1 - 0.96\theta_i)/(1 + 0.5\theta_i)], \quad (12)$$

where θ_i is the local volume fraction of counterions (including frozen water molecules),

$$\theta_i = c_i[\mathrm{Na}^+]/(c_i[\mathrm{Na}^+] + 11.2). \quad (13)$$

$c_i[\mathrm{Na}^+]$ is the local molar concentration of counterions in element i, and ϵ_i^{BE}, the local dielectric constant due to boundary and electric-field effects. Equation (12) is used in Eq. (4) and the charge distributions [Eq. (3)] and potentials [Eq. (4)] are iterated until self-consistency. The result was considered converged when the sum of the average rms values of the total charges, potentials, and field strengths at all environmental points between consecutive iterations is less than 10^{-6}.

Results and Conclusions

The polyelectrolyte system that we consider is identical to that used in [16]. An all-atom model of B-DNA was contained within a cylindrical cell of 100 Å radius with added salt in concentrations of 0, 0.1, 0.2, and 0.5M. As before, environmental points were placed in layers contouring the van der Waals' surface of the macromolecule. The data in Tables I and II are average potentials and cation concentrations by groove for calculations *with* [Eq. (12)] and *without* ($\epsilon_i = 78.5$) the inclusion of the dielectric effects discussed above. The results for average local dielectric constants in the grooves is given in Table III. In all cases, the dielectric constant of cells partitioning the DNA was set to 4 and a temperature of 298 K assumed. In their closest approach to the surface of DNA, solvated sodium counterions were modeled as 1.4 Å radius spheres.

Table I shows that the effect of including a variable dielectric constant upon the electrostatic potential (relative to that at the cell boundary) is confined mainly to the first environment layer at the surface of DNA. While limited in radial extent, the deviations range from the fixed dielectric constant potentials by 50–100%. The size of the effect increases with added salt, due, of course, to an increase in surface counterion concentrations and accounted for by Eq. (12). The effect is only moderately sensitive to angular placement around the surface of the DNA.

While counterions are absent from this inner Helmholtz layer, the increase in surface potential may be relevant to the calculation of pK_a values at specific protonation sites. Local concentrations of solvated protons are readily calculated from the potentials given in Table I through $\mathrm{p[H]}_i = \mathrm{pH} + 0.017\phi_i$ (mV) [6]. For a pH value of 7 at the cell boundary ($\phi = 0$), average p[H] values in the second layer (2.1 Å from the surface) obtained by including local dielectric effects are only about 0.1 unit lower than without the effects giving values of 4.4, 5.6, 5.8, and 6.1 at added salt concentrations of 0, 0.1, 0.2, and 0.5M, respectively, with angular deviations less than 0.2 units. However, the much steeper potential gradient at the surface in the presence of local dielectric effects implies that a reliable distance of the closest approach to the DNA surface for the proton (as a hydronium ion)

TABLE I

Average potentials (mV, relative to the potential at the cell boundary) in the major and minor grooves and elsewhere of B-DNA in the presence of various concentrations of added salt with [Eq. (12)] and without ($\epsilon = 78.5$) the inclusion of the algorithm for local dielectric constants; distances are measured relative to the van der Waals surface of DNA.

		Potential (mV)							
		Fixed	solvent	dielectric	constant	Variable	solvent	dielectric	constant
R (Å)		0*M*	0.1*M*	0.2*M*	0.5*M*	0*M*	0.1*M*	0.2*M*	0.5*M*
0.7	Major groove	−205	−139	−124	−105	−319	−255	−241	−225
	Minor groove	−145	−77	−62	−42	−177	−109	−93	−73
	Elsewhere	−173	−107	−92	−74	−254	−190	−176	−160
	Average	−187	−121	−106	−87	−280	−215	−202	−186
2.1	Major groove	−149	−83	−68	−50	−152	−86	−71	−52
	Minor groove	−141	−74	−59	−40	−157	−88	−72	−52
	Elsewhere	−141	−76	−61	−43	−149	−83	−69	−50
	Average	−145	−79	−64	−46	−151	−85	−70	−51
3.5	Major groove	−124	−60	−46	−30	−120	−57	−43	−27
	Minor groove	−127	−62	−47	−30	−129	−63	−48	−31
	Elsewhere	−123	−59	−45	−28	−122	−58	−44	−28
	Average	−124	−60	−46	−29	−123	−58	−44	−28
5.9	Average	−98	−37	−25	−13	−93	−34	−23	−11
9.3	Average	−79	−24	−14	−5	−75	−21	−12	−4
13.0	Average	−62	−13	−7	−2	−59	−12	−6	−1

TABLE II

Average cation concentrations (*M*) near B-DNA as described in Table I.

		Concentration (M)							
		Fixed	solvent	dielectric	constant	Variable	solvent	dielectric	constant
R (Å)		0*M*	0.1*M*	0.2*M*	0.5*M*	0*M*	0.1*M*	0.2*M*	0.5*M*
2.1	Major groove	3.12	3.53	3.76	4.27	3.16	3.59	3.82	4.33
	Minor groove	1.66	1.93	2.08	2.44	3.04	3.37	3.55	3.95
	Elsewhere	1.82	2.16	2.36	2.84	2.64	3.06	3.30	3.84
	Average	2.38	2.74	2.95	3.43	2.91	3.33	3.56	4.07
3.5	Major groove	0.92	1.17	1.32	1.67	0.75	0.99	1.13	1.47
	Minor groove	0.97	1.19	1.32	1.64	1.04	1.26	1.39	1.72
	Elsewhere	0.88	1.09	1.23	1.55	0.86	1.07	1.20	1.52
	Average	0.91	1.13	1.27	1.61	0.85	1.07	1.20	1.53
5.9	Average	0.32	0.46	0.56	0.84	0.26	0.41	0.50	0.78
9.3	Average	0.15	0.27	0.36	0.63	0.13	0.24	0.33	0.60
13.0	Average	0.07	0.18	0.27	0.54	0.06	0.17	0.26	0.54

TABLE III
Average local dielectric constants in the major and minor grooves and elsewhere of B-DNA in the presence of various concentrations of added salt calculated using Eq. (12) in the Poisson – Boltzmann equation.

		Dielectric constant			
R (Å)		0*M*	0.1*M*	0.2*M*	0.5*M*
0.7	Major groove	20	20	20	20
	Minor groove	46	46	45	45
	Elsewhere	23	23	23	23
	Average	23	23	23	23
2.1	Major groove	31	29	29	27
	Minor groove	29	28	27	26
	Elsewhere	34	33	32	30
	Average	32	31	30	28
3.5	Major groove	59	58	57	56
	Minor groove	41	40	39	38
	Elsewhere	55	54	53	52
	Average	54	53	53	51
5.9	Average	71	71	70	69
9.3	Average	76	76	75	73
13.0	Average	77	77	76	73

is crucial. Assuming that this distance is 1.4 Å leads to the above p[H] values. It is more likely that solvated protons can approach the surface closer than this. If a distance of 1 Å is assumed, surface proton concentrations are then much more sensitive to local dielectric effects as well as showing a much greater angular dependence. Approximate local p[H] values may be (quadratically) interpolated from the potentials at 1.5 Å from the surface to give values in the major groove of 3.5, 4.6, 4.9, and 5.2 at added salt concentrations of 0, 0.1, 0.2, and 0.5*M*, respectively. The values average 0.65 units lower than when a constant solvent dielectric constant is used. Minor groove p[H] values average 0.75 units higher than major groove values, an observation in the opposite direction of that found previously [6]. This is due to the specific sizes and concentrations of counterions assumed in the two studies as well as to the incorporation of activity corrections to the standard PB approach. This emphasizes the importance of local surface-counterion geometries and conditions in determining specific thermodynamic properties, such as local p[H], which, for DNA, may provoke a selective biochemical response [24]. If future PB applications are to focus on surface-determined properties of polyelectrolytes, as indeed recent work on pK_a's seems to indicate, then a much better analysis and improvement of the PB theory at such surfaces is warranted.

Counterion concentrations consistent with the potentials of Table I are given in Table II. The largest deviations are observed in the counterion layer closest to the surface, which corresponds to the second layer listed in Table I. Of particular interest is the sensitivity of minor groove counterion concentrations to variable dielectric effects as well as the constancy of major groove concentrations. With counterion concentrations already saturating the major groove, inclusion of a variable local dielectric constant does not increase values there much. However, elsewhere at the surface—particularly within the minor groove—concentrations increase toward major groove values. In fact, in the second ion layer, the minor groove concentration is larger than the major groove value. Isolation of the individual dielectric decrement effects shows that the electric-field component plays a much larger role in the minor groove than in the major groove.

From Table III, the average solvent dielectric constant in the grooves is seen to be essentially independent of added salt concentration. With the exception of the minor groove, the local dielectric constant of all other regions monotonically increases away from the surface toward the bulk value. The minor groove, however, displays a minimum dielectric constant about 2 Å from the surface due to the presence of counterions as well as the high electric-field strength in the region.

Whereas potential and counterion concentrations are affected mainly in the layer closest to the surface, solvent dielectric constant values are still 10% below the bulk value as far as 6 Å from the surface. The significance of this in accurately modeling the electrostatic force between a molecule (protein, ligand) and DNA or between an electron donor–acceptor pair interacting near the surface is clear.

However, for processes further from the surface or for those where an average over the entire environment space is required, the lowering of the dielectric constant at the DNA surface may not be as important. Consider the charged cylinder cell model calculation of our previous work [16] in which 50 mM monovalent salt was added to a neutral system consisting of a 10 Å radius cylinder with the same surface charge density as DNA in a cell of 100 Å radius. The first passage time for an ion or molecular dipole to diffuse from the outer cell to the cylinder surface can easily be calculated given the interaction energy (involving the PB electrostatic potential, the ionic charge or dipole moment, and the spatially varying dielectric constant) as a function of the radial distance from the cylinder [25]. Assuming that the diffusion constant for the ion or dipole is spatially invariant, the first passage time with a variable solvent dielectric constant for mono- or divalent charges is only about 2% less than that with a fixed solvent dielectric constant equal to 78.5. For dipole moments of 2 and 10 D, the deviations are only 1 and 3% less, respectively. These results are not totally unexpected since the first passage time is not strongly dependent on the interaction energy.

The results presented here are only suggestive of what is likely to be observed in other systems. Ion size, surface geometry, electrostatic potential, counterion concentration, and local dielectric constants are all closely coupled so that only broad generalities may be drawn. It is likely that certain cases would warrant the inclusion of local dielectric constants above others with the particular thermodynamic property being investigated of prime consideration. In those cases, the effect may be considerable.

ACKNOWLEDGMENT

This work was supported by Grant GM29079 from the National Institutes of Health.

References

1. C. F. Anderson and M. T. Record, Jr., Annu. Rev. Biophys. Chem. **19**, 423 (1990).
2. H. C. Andersen and D. Chandler, J. Chem. Phys. **57**, 1918 (1972).
3. P. Mills, C. F. Anderson, and M. T. Record, Jr., J. Chem. Phys. **89**, 3984 (1985).
4. C. S. Murthy, R. J. Bacquet, and P. J. Rossky, J. Phys. Chem. **82**, 701 (1985).
5. G. R. Pack, G. Lamm, L. Wong, and D. Clifton, in *Theoretical Biochemistry and Molecular Biophysics*, D. L. Beveridge and R. Lavery, Eds. (Adenine Press, New York, 1990), pp. 237–246.
6. G. Lamm and G. R. Pack, Proc. Natl. Acad. Sci. U.S.A. **87**, 9033 (1990).
7. G. Lamm and G. R. Pack, unpublished.
8. T. Das, D. Bratko, L. B. Bhuiyan, and C. W. Outhwaite, J. Phys. Chem. **99**, 410 (1995).
9. J. Warwicker and H. C. Watson, J. Mol. Biol. **157**, 671 (1982).
10. J. Warwicker, J. Theor. Biol. **121**, 199 (1986).
11. J. Warwicker, J. Mol. Biol. **236**, 887 (1994).
12. I. Klapper, R. Hagstrom, R. Fine, K. Sharp, and B. Honig, Proteins **1**, 47 (1986).
13. M. K. Gilson, A. Rashin, R. Fine, and B. Honig, J. Mol. Biol. **183**, 503 (1985).
14. M. K. Gilson, M. E. Davis, B. A. Luty, and J. A. McCammon, J. Phys. Chem. **397**, 3591 (1993).
15. G. R. Pack, G. A. Garrett, L. Wong, and G. Lamm, Biophys. J. **365**, 1363 (1993).
16. G. Lamm and G. R. Pack, J. Phys. Chem B **101**, 959 (1997).
17. See, e.g., T. L. Hill, *An Introduction to Statistical Thermodynamics* (Dover, New York, 1986), Chap. 18.
18. J. G. Kirkwood, J. Chem. Phys. **7**, 911 (1939).
19. H. Fröhlich, *Theory of Dielectrics*, 2nd ed. (Clarendon Press, Oxford, 1958).
20. F. Booth, J. Chem. Phys. **19**, 391, 1327, 1615 (1951).
21. R. Pottel, in *Water: A Comprehensive Treatise*, F. Franks, Ed. (Plenum Press, New York, 1973), Vol. 3, Chap. 8.
22. G. H. Haggis, J. B. Hasted, and T. J. Buchanan, J. Chem. Phys. **20**, 1452 (1952).
23. J. B. Hasted, in *Water: Comprehensive Treatise*, F. Franks, Ed. (Plenum Press, New York, 1973), Vol. 3, Chap. 7.
24. G. Lamm, L. Wong, and G. R. Pack, J. Am. Chem. Soc. **118**, 3325 (1996).
25. A. Szabo, K. Schulten, and Z. Schulten, J. Chem. Phys. **72**, 4350 (1980).

Valence Ionization Potentials of Anionic Phosphate Esters: An Ab Initio Quantum Mechanical Study

SHARON M. FETZER,[1] PIERRE R. LEBRETON,[1] MARIE-MADELAINE ROHMER,[2] ALAIN VEILLARD[2]

[1] *Department of Chemistry, University of Illinois at Chicago, Chicago, Illinois 60607*
[2] *UPR 139 du CNRS, Institut Le Bel–4, rue Blaise Pascal, 67000 Strasbourg, France*

Received 5 March 1997; accepted 6 June 1997

ABSTRACT: In earlier investigations N. S. Kim, P. R. LeBreton, J. Am. Chem. Soc. **118**, 3694 (1996), and references therein, ultraviolet (UV) photoelectron data and ab initio molecular orbital calculations yielded information about nucleotide electronic structure by providing valence ionization potentials (IPs) of nucleotide bases and sugar model compounds. Here, model phosphate group ionization potentials have been evaluated by employing multireference, single and double excitation configuration interaction calculations with a complete active space self-consistent field (CASSCF) wave function. The five lowest energy IPs of $H_2PO_4^-$ and the four lowest energy IPs of $(CH_3)HPO_4^-$ and $(CH_3)_2PO_4^-$ were evaluated. Calculations were performed using a $(12,9,1)/[6,4,1]$ double-zeta basis set on phosphorus with a *d*-polarization function; $(10,7)/[4,3]$ and $(10,6)/[4,2]$ basis sets on oxygen and carbon, respectively; and a $(6)/[3]$ basis set on hydrogen. Two types of CASSCF calculations were carried out. In one, denoted $8e8a/7e8a$, the anions and radicals had 8 and 7 electrons, respectively, in 8 active orbitals. In the second, denoted $10e10a/9e10a$, there were 9 and 10 electrons in 10 active orbitals. Ionization potentials of $H_2PO_4^-$, $(CH_3)HPO_4^-$, and $(CH_3)_2PO_4^-$ were also obtained from second-order perturbation theory (CASPT2) calculations with the CASSCF reference functions. All of the ionization events examined are associated with removal of electrons from oxygen atom lone-pair orbitals on the closed-shell anions. For $H_2PO_4^-$ and $(CH_3)HPO_4^-$, the lowest energy CASPT 2 ionization potentials obtained (4.29–4.36 eV and 4.12–4.27 eV, respectively) are smaller than corresponding IPs (4.89 and 4.69 eV) previously reported from results of 6-31 + G* second-order Møller-Plesset (MP2) calculations. For the second through fifth IPs of $H_2PO_4^-$, and the second through fourth IPs of $(CH_3)HPO_4^-$, the values obtained from CASPT2 calculations are 0.42 to 1.95 eV smaller than values reported from the combined use of MP2 and configuration interaction

Correspondence to: P. R. LeBreton.
Contract grant sponsor: American Cancer Society.
Contract grant number: CN-37E.

International Journal of Quantum Chemistry, Vol. 65, 1095–1106 (1997)
© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061095-12

singles calculations with the 6-31 + G* basis set. A combination of results from MP2 and CASPT2 calculations yields values of 4.89, 5.42, 6.30, 6.64, and 7.41 eV for the five lowest IPs of $H_2PO_4^-$; and values of 4.69, 5.43, 6.08, and 6.55 eV for the four lowest IPs of $(CH_3)HPO_4^-$. © 1997 John Wiley & Sons, Inc. Int J Quant Chem 65: 1095–1106, 1997

Introduction

Recent quantum mechanical calculations have provided ionization potentials [1–4], electron affinities [3, 4], and vibration [5, 6] transition energies in nucleotide components. Recent results from quantum calculations have also described electronic influences on tautomerism [5–7], hydrogen bonding and solvation [8], base stacking [9], proton transfer in radical anions [10], radiation-induced modification mechanisms [11], and alkylation reactions [12–16] involving nucleotides, nucleotide components, and nucleotide model compounds. Essential to all of these investigations are reliable and accurate descriptions of valence electrons.

Because nucleotides contain between 160 and 180 electrons, it is not currently possible to describe their valence structure by employing highly rigorous computational methods. To date, the electronic structures of intact nucleotides have only been examined at the self-consistent field (SCF) level [9, 12, 13, 17–21]. The results of these calculations indicate that the upper occupied π and lone-pair orbitals are largely localized on either the base, sugar, or phosphate groups and have electron distributions which are similar to those occurring in the separated components. While it is not possible to easily assess the reliability of SCF descriptions of valence electrons in intact nucleotides, it is possible to test descriptions of nucleotide components. One strategy, employed by two of the contributors to this investigation (S. M. F. and P. R. L.), has been to compare calculated ionization potentials (IPs) of nucleotide base and sugar model compounds, obtained by applying Koopmans' theorem [22] to SCF results, with experimental IPs obtained from He(I) ultraviolet (UV) photoelectron (PE) experiments [1, 2, 12, 13, 23, 24]. In addition to testing the SCF results for nucleotide components, these comparisons of theoretical and experimental IPs have provided a means by which valence IPs obtained from SCF calculations on intact nucleotides can be individually corrected so as to provide more reliable values of as many as 14 of the lowest energy IPs in the nucleotides [9, 12, 13, 18–21].

While PE data for nucleotide bases and sugar model compounds have provided an experimental basis for assessing theoretical descriptions of the valence electrons in the neutral base and sugar groups of nucleotides, experimental IPs are available for only a few oxygen and phosphorus containing anions. It is necessary to obtain an understanding of phosphate esters which is as reliable as the current understanding of bases and sugars in order to develop highly accurate descriptions of the electronic properties of nucleotides. However, to our knowledge, no experimental IPs have yet been reported for anionic phosphate esters.

In earlier theoretical investigations, SCF calculations on phosphate esters were employed to examine orbital energies [11, 25], electrostatic potentials [26], and conformational properties [27]. In post-SCF investigations [12, 13], second-order (MP2) and third-order (MP3) Møller-Plesset perturbation calculations with the 6-31 + G* basis set [28] (MP2/ and MP3/6-31 + G*) yielded values of 4.89 and 4.87 eV, respectively, for the first ionization potential of $H_2PO_4^-$. Here, the first IP was obtained as the difference between the energies of the ground-state, closed-shell anion and the ground-state radical formed by removal of an electron. In test MP2/6-31 + G* calculations on CH_3O^-, PO_2^-, and PO_3^-, and the corresponding ground-state radicals, theoretical IPs obtained in this manner differed from experimental IPs by less than 0.3 eV [12]. In an extension of this approach, excitation energies of $H_2PO_4^{\cdot}$ obtained from 6-31 + G* calculations using the configuration interaction singles (CIS) method [29], were combined with results from MP2 calculations on the ground-states of $H_2PO_4^-$ and $H_2PO_4^{\cdot}$. This combination of calculations (denoted MP2/CIS) provided values for the second to fifth lowest energy IPs of $H_2PO_4^-$. The same approach was also applied to $(CH_3)HPO_4^-$ [18]. When employed to calculate the first five IPs of NO_2, the MP2/CIS method yielded a value of the lowest energy IP which differed from the experimental value by less

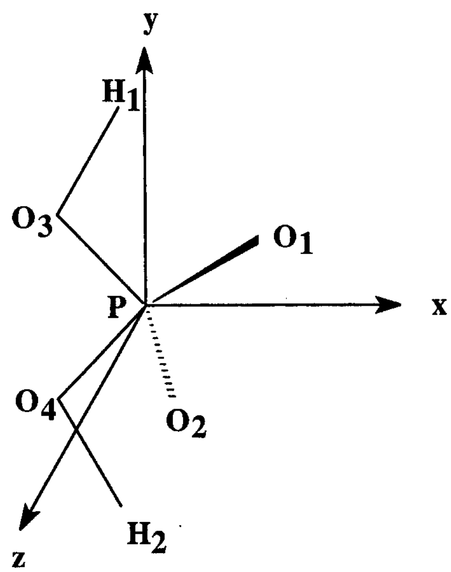


FIGURE 1. Choice of axis for $H_2PO_4^-$.

than 0.1 eV. However, the theoretical values of the higher energy IPs were less accurate. In one case, the experimental IP differed from the theoretical IP by more than 0.9 eV [12].

The goal of the present investigation is to employ a more rigorous computational approach which relies on the multireference CI method with singles and doubles, and a complete active space self-consistent field (CASSCF) wave function [30]. Second-order pertubation theory (CASPT2) calculations with the CASSCF wave functions [31, 32] have also been employed to describe the five lowest energy ionization events in $H_2PO_4^-$, and the four lowest energy ionization events ionization events in $CH_3HPO_4^-$ and $(CH_3)_2PO_4^-$.

Methods

The symmetry of $H_2PO_4^-$ and $(CH_3)_2PO_4^-$ employed in the calculations was C_{2v}; the symmetry of $CH_3HPO_4^-$ was C_s. The geometry for $H_2PO_4^-$ has been obtained previously by optimization at the 6-31 + G* level [12, 28]. The choice of axis for $H_2PO_4^-$ is shown in Figure 1. For $CH_3HPO_4^-$, heavy-atom bond lengths were obtained from deoxcytidylyl(3′-5′) deoxyguanosine crystallographic data [18, 33]. The geometry used for $(CH_3)_2PO_4^-$ was based on those for $H_2PO_4^-$ and $(CH_3)HPO_4^-$. Bond lengths and bond angles used for calculations on the phosphate anions are given in Table I. The geometries of the radicals were taken to be the

TABLE I
Bond lengths and bond angles of $H_2PO_4^-$, $CH_3HPO_4^-$, and $(CH_3)_2PO_4^-$.

Bond	Length[a]	Atoms	Angle[b]
P—O	1.63	O=P—O	109.74
P=O	1.48	O—P—O	94.74
O—H	0.95	P—O—H	109.13
O—C	1.40	P—O—C	117.69
C—H	1.08	O—C—H	110.96

[a] In angstroms
[b] In degrees.

same as those of the anions, and the calculated IPs correspond to vertical ionization potentials.

Polarization and diffuse functions are needed to obtain reliable results on phosphate anions [25, 34, 35]. For phosphorus a (12, 9, 1) Gaussian basis set was contracted to [6, 4, 1], resulting in a double-zeta basis set [37] augmented with one d-polarization function, with an orbital exponent $[\alpha_d(P)]$ of 0.43 [38]. Values of the $3d$ exponent, found in the literature, range between 0.39 and 0.60 [25, 28, 36, 38–41]. An oxygen (10, 7) basis set, contracted to [4, 3], was derived from a previously reported (10, 6) basis set [42] by adding one set of diffuse p functions. The value of the orbital exponent (α_p) was 0.0564 and was chosen according to the even-temperted criterion. In comparison, the value of the exponent optimized for the oxygen anion is 0.0515. For carbon and hydrogen, previously reported (10, 6) and (6) basis sets [42, 43] were contracted to [4, 2] and [3], respectively.

Ionization potentials were obtained as the difference between the energies calculated for the anion and the radial at the SCF, CASSCF, and CASPT2 levels. For $H_2PO_4^-$, one set of IPs was obtained from a CASSCF calculation on $H_2PO_4^-$ (denoted $8e8a$) in which there were eight electrons in eight active orbitals, two within each irreducible reprentation. In conjunction with the $8e8a$ calculation on $H_2PO_4^-$, a CASSCF calculation (denoted $7e8a$) was also carried out on the $H_2PO_4^{\cdot}$ radical. This had seven electrons in eight active orbitals. The $H_2PO_4^-$ and $H_2PO_4^{\cdot}$ calculations yielded values (denoted $8e8a/7e8a$) for the four lowest ionization potentials. The four IPs corresponded, respectively, to the removal of one electron from the highest occupied molecular orbital of each irreducible representation (orbitals $11a_1$, $2a_2$, $6b_1$, and $6b_2$). For $H_2PO_4^-$, a second set of ionization poten-

tials (denoted $10e10a/9e10a$) were obtained from $10e10a$ calculations on $H_2PO_4^-$, and $9e10a$ calculations on $H_2PO_4^{\cdot}$. Of the 10 active orbitals, 4 have b_1 symmetry, and 2 each have a_1, a_2, and b_2 symmetries. This set of calculations yielded the five lowest energy IPs of $H_2PO_4^-$. The fifth IP corresponds to removal of an electron from the $5b_1$ orbital. For $H_2PO_4^-$, the first and fifth IPs, which both arise from removal of a b_1 electron were obtained via a third set of CASSCF calculations. In this case, the ionization potentials (denoted $10e10a$/av.$9e10a$) were obtained from calculations for which the average energies associated with the two $^2B_1$ states of $H_2PO_4^{\cdot}$ (a^2B_1 and b^2B_1) were optimized.

For $CH_3HPO_4^-$, ionization potentials were also obtained from $8e8a/7e8a$ and $10e10a/9e10a$ calculations. However, the CASSCF calculation of the lowest roots for the $^2A'$ and $^2A''$ states of $CH_3HPO_4^{\cdot}$ yields only two IPs, corresponding to revmoal of an electron from the $20a'$ and $9a''$ orbitals, respectively. Attempts to perform CASSCF calculations with the optimization carried out for the second root corresponding to the $^2A'$ and $^2A''$ states were unsuccessful. In order to obtain IPs associated with the two lowest energy $^2A'$ states of $CH_3HPO_4^{\cdot}$, and the two lowest energy $^2A''$ states, we performed CASSCF calculations optimized for the average energy of the two lowest states of each symmetry. These states, for which the average energies were optimized, are denoted a^2A'' and b^2A'', and a^2A' and b^2A', respectively. Values of IPs, based on the calculations for which the average energies were optimized, are denoted $8e8a$/av.$7e8a$ and $10e10a$/av.$9e10a$. For $(CH_3)_2PO_4^-$, values of the four lowest energy IPs were obtained from $8e8a/7e8a$ calculations.

Second-order perturbation theory (CASPT2) calculations with the reference CASSCF $8e8a/7e8a$, $8e8a/av.7e8a$, $10e10a/9e10a$ and $10e10a$/av.$9e10a$ wave functions have also been employed to evaluate selected IPs of $H_2PO_4^-$, $CH_3HPO_4^-$, and $(CH_3)_2PO_4^-$. The CASPT2 energy values reported correspond to the nondiagonal approach. Finally, the first five IPs of $H_2PO_4^-$ and the first four IPs of $CH_3HPO_4^-$ have been obtained by combining previously reported [12, 18, 19] values for the lowest energy IPs obtained from MP2/6-31 + G* calculations with results from CASPT2 calculations.

Results and Discussion

The total energies of all species are listed in Tables II to IV, and the corresponding IPs are reported in Table V. In Table V, IPs calculated at the CASSCF level are 0.28–0.81 eV smaller than IPs calculated at the SCF level. This difference arises because the closed-shell anion is better represented by a single determinant wave function than the open-shell radicals. For $H_2PO_4^-$, the coef-

TABLE II
Total energies of $H_2PO_4^-$ and $H_2PO_4^{\cdot}$ in a.u.[a]

	$H_2PO_4^-$	$H_2PO_4^{\cdot}$				
	$^1A_1$	a^2B_1	$^2A_2$	$^2B_2$	$^2A_1$	b^2B_1
E(SCF)	−0.477068	−0.305359	−0.277967	−0.237350	−0.227009	
E(CASSCF)[b]	−0.573979	−0.418271	−0.402465	−0.359464	−0.342319	
E(CASSCF)[c]	−0.592194	−0.438787	−0.418125	−0.375717	−0.371143	−0.322455
E(CASSCF)[d]		−0.430378				−0.286281
E(CASPT2)[e]	−0.149818[f]	−0.989762	−0.971146	−0.939456	−0.925549	
E(CASPT2)[g]	−0.151389[f]	−0.992279	−0.972808	−0.940458	−0.927958	−0.899667
E(CASPT2)[h]		−0.993831				−0.914097

[a] Relative to −641.000000 a.u. unless otherwise indicated.
[b] CASSCF $8e8a$ or $7e8a$.
[c] CASSCF $10e10a$ or $9e10a$.
[d] CASSCF $9e10a$ optimized for the average energy of the a^2B_1 and b^2B_1 states of $H_2PO_4^{\cdot}$ (denoted CASSCF av. $9e10a$).
[e] With the CASSCF $8e8a$ or $7e8a$ reference wave functions.
[f] Relative to −642.000000 a.u.
[g] With the CASSCF $10e10a$ or $9e10a$ reference wave functions.
[h] With the CASSCF av. $9e10a$ reference wave function.

TABLE III
Total energies of $CH_3HPO_4^-$ and $CH_3HPO_4^{\cdot}$ in a.u.[a]

	$CH_3HPO_4^-$	$CH_3HPO_4^{\cdot}$			
	$^1A'$	a^2A''	b^2A''	a^2A'	b^2A'
E(SCF)	−0.482995	−0.314723		−0.248211	
E(CASSCF)[b]	−0.579359	−0.428128		−0.368744	
E(CASSCF)[c]		−0.426110	−0.402247	−0.357785	−0.350327
E(CASSCF)[d]	−0.599325	−0.450873		−0.394395	
E(CASSCF)[e]		−0.449348	−0.422732	−0.380919	−0.372202
E(CASPT2)[f]	−0.249707	−0.092824		−0.042585	
E(CASPT2)[g]		−0.092638		−0.041766	−0.028106
E(CASPT2)[h]	−0.250381	−0.093604		−0.046247	
E(CASPT2)[i]		−0.098937	−0.071913		−0.030707

[a] SCF and CASSCF energies are relative to −680.000000 a.u. CASPT2 energies are relative to −681.000000 au.
[b] CASSCF 8*e*8*a* or 7*e*8*a*.
[c] CASSCF 7*e*8*a* optimized for the average energy of the two a^2A' and b^2A' states, and two a^2A'' and b^2A'' states (denoted CASSCF av.7*e*8*a*).
[d] CASSCF 10*e*10*a* or 9*e*10*a*.
[e] CASSCF 9*e*10*a* optimized for the average energy of the two a^2A' and b^2A' states, and the two a^2A'' and b^2A'' states (denoted CASSCF av.9*e*10*a*).
[f] With the CASSCF 8*e*8*a* or 7*e*8*a* refrence wave functions.
[g] With the CASSCF av.7*e*8*a* reference wave function.
[h] With the CASSCF 10*e*10*a* or 9*e*10*a* reference wave functions.
[i] With the CASSCF av.9*e*10*a* reference wave function.

ficient of the leading configuration in the CASSCF wave function is 0.975. For the ground-state radicals, the coefficients of the leading configuration are in the range 0.95–0.96.

The results in Table V demonstrate that ionziation potentials from a CASSCF calculation optimized for the average energy of two states are not accurate. When calculations were carried out with a 9*e*10*a* reference wave function, optimized for the average energy of the a^2B_1 and b^2B_1 states of $H_2PO_4^{\cdot}$, the CASSCF 10*e*10*a*/av.9*e*10*a* values for the a^2B_1 and b^2B_1 ionization potentials were 4.40 and 8.32 eV. When the calculation was optimized for the a^2B_1 and b^2B_1 states separately, the corresponding CASSCF 10*e*10*a*/9*e*10*a* ionization potentials were 4.17 and 7.34 eV. In contrast, the IPs obtained from CASPT2 calculations employing an averaged CASSCF 9*e*10*a* reference wave function to describe $H_2PO_4^{\cdot}$ (CASPT2 10*e*10*a*/av.9*e*10*a*) are in closer agreement with IPs obtained from CASPT2 calculations with separately optimized wave functions (CASPT2 10*e*10*a*/9*e*10*a*). The CASPT2 10*e*10*a*/av.9*e*10*a* values for the a^2B_1 and b^2B_1 ionization potentials of $H_2PO_4^-$ are 4.29 and 6.46 eV, respectively; the CASPT2 10*e*10*a*/9*e*10*a* values are 4.33 and 6.85 eV. Similar trends have

TABLE IV
Total energies of $(CH_3)_2PO_4^-$ and $(CH_3)_2PO_4^{\cdot}$ in a.u.[a]

	$(CH_3)_2PO_4^-$	$(CH_3)_2PO_4^{\cdot}$			
	$^1A_1$	$^2B_1$	$^2A_2$	$^2B_2$	$^2A_1$
E(SCF)	−0.488963	−0.324054	−0.294497	−0.257729	−0.248975
E(CASSCF)[b]	−0.585302	−0.435846	−0.418264	−0.377721	−0.361671
E(CASPT2)[c]	−0.349900	−0.195698	−0.173287	−0.147411	−0.134618

[a] SCF and CASSCF energies relative to −719.000000 a.u. CASPT2 energies relative to −720.000000 a.u.
[b] CASSCF 8*e*8*a* or 7*e*8*a*.
[c] With the CASSCF 8*e*8*a* or 7*e*8*a* reference wave functions.

TABLE V
Calculated ionization potentials (in eV) of $H_2PO_4^-$, $CH_3HPO_4^-$, and $(CH_3)_2PO_4^-$.

	Anions												
	$H_2PO_4^-$					$CH_3HPO_4^-$				$(CH_3)_2PO_4^-$			
	Radical States												
	a^2B_1	$^2A_2$	$^2B_2$	$^2A_1$	b^2B_1	a^2A''	b^2A''	a^2A'	b^2A'	$^2B_1$	$^2A_2$	$^2B_2$	$^2A_1$
SCF	4.68	5.41	6.53	6.80		4.57		6.39		4.49	5.28	6.29	6.50
CASSCF[a]	4.22	4.67	5.84	6.30		4.12		5.74		4.07	4.54	5.65	6.08
CASSCF[b]	4.17	4.74	5.89	6.01	7.34	4.04		5.58					
CASSCF[c]						4.18	4.82	6.03	6.24				
CASSCF[d]	4.40				8.32	4.08	4.80	5.94	6.18				
CASPT2[e]	4.36	4.86	5.72	6.10		4.27		5.63		4.20	4.81	5.51	5.86
CASPT2[f]	4.33	4.86	5.74	6.08	6.85	4.27		5.55					
CASPT2[g]						4.27		5.66	6.03				
CASPT2[h]	4.29				6.46	4.12	4.86		5.98				

[a] CASSCF 8*e*8*a* / 7*e*8*a*.
[b] CASSCF 10*e*10*a* / 9*e*10*a*.
[c] CASSCF 8*e*8*a* / av.7*e*8*a*.
[d] CASSCF 10*e*10*a* / av.9*e*10*a*.
[e] CASPT2 8*e*8*a* / 7*e*8*a*.
[f] CASPT2 10*e*10*a* / 9*e*10*a*.
[g] CASPT2 8*e*8*a* / av.7*e*8*a*.
[h] CASPT2 10*e*10*a* / av.9*e*10*a*.

been observed in other calculations of exicted states [44].

The most rigorously calculated first ionization potentials of $H_2PO_4^-$ and $CH_3HPO_4^-$ were obtained from CASPT2 $10e10a/9e10a$ calculations and have values of 4.33 and 4.27 eV, respectively. A comparison of CASPT2 results for $H_2PO_4^-$ and $CH_3HPO_4^-$ indicates that for each anion the first IPs obtained with the $9e10a$ and av.$9e10a$ $H_2PO_4^{\cdot}$ reference wave functions differ by no more than 0.15 eV. For the higher energy IPs, larger differences sometimes occur. For example, for the fifth IP of $H_2PO_4^-$, and the second IP of $CH_3HPO_4^-$, the differences between the CASPT2 values obtained with the $9e10a$ and the av.$9e10a$ reference wave functions are 0.39 and 0.43 eV, respectively.

The CASPT2 $10e10a/9e10a$ values of the first four IPs of $H_2PO_4^-$, and of the first and third IPs of $CH_3HPO_4^-$ differ from the CASPT2 $8e8a/7e8a$ values by no more than 0.08 eV. The CASPT2 $10e10a/9e10a$ values of the fist IPs of $H_2PO_4^-$ and $CH_3HPO_4^-$ (4.33 and 4.27 eV) are 0.56 and 0.42 eV smaller than IPs obtained from MP2/6-31 + G* calculations [12, 18, 19]. In test MP2/6-31 + G* calculations on CH_3O^-, PO_2^-, and PO_3^-, theoretical first IPs differed from experimental values by less than 0.3 eV [12]. This observation suggests that the first IPs predicted by the MP2/6-31 + G* calculations are more accurate than the first IPs obtained from the CASPT2 calculations. Of course, the MP2/6-31 + G* calculations provide no reliable information about higher energy ionization potentials.

With the 6-31 + G* basis set, earlier investigations [12, 18] employing a combination of MP2 and CIS methods (denoted MP2/CIS) provided values for the second through fifth IPs of $H_2PO_4^-$, associated with the $^2A_2$, $^2B_2$, $^2A_1$, and b^2B_1 states of $H_2PO_4^{\cdot}$. Similarly, MP2/CIS calculations [18, 19] were carried out to obtain the second through fifth IPs of $CH_3HPO_4^-$, which give rise to the a^2A'', b^2A'', a^2A', and b^2A' states of $CH_3HPO_4^{\cdot}$. The CIS excitation energies for $H_2PO_4^{\cdot}$ are, in some cases, significantly different from CASPT2 and CASSCF exciation energies. For $H_2PO_4^{\cdot}$, the first ($^2A_2$), second ($^2B_2$) and fourth (b^2B_1) excitation energies, and for $CH_3HPO_4^{\cdot}$, the first (b^2A'') and second (a^2A') excitation energies obtained from all of the CASPT2 calculations are 0.47 to 1.74 eV smaller than corresponding excitation energies obtained from CIS calculations. In cases where comparisons were made, results from CASSCF calculations were similar. Here, the CASSCF $7e8a$ and $9e10a$ excitation energies for the $^2A_2$, $^2B_2$, and b^2B_1 states of

$H_2PO_4^{\cdot}$, and for the a^2A' state of $CH_3HPO_4^{\cdot}$ are between 0.52 and 0.87 eV smaller than the CIS energies.

The CASPT2 and CASSCF values for the third excitation energies associated with the $^2A_1$ and b^2A' states of $H_2PO_4^{\cdot}$ and $CH_3HPO_4^{\cdot}$, respectively, are larger than values obtained from CIS calculations. Results from CASPT2 $7e8a$ and $9e10a$ calculations on $H_2PO_4^{\cdot}$ and av.$7e8a$ and av. $9e10a$ calculations on $CH_3HPO_4^{\cdot}$ yield $^2A_1$ and b^2A' excitation energies that are 0.27–0.40 eV larger than the CIS energies. Similar results were obtained from CASSCF $7e8a$ and $9e10a$ calculations on $H_2PO_4^{\cdot}$, which predict that the $^2A_1$ excitation energy is 0.37–0.61 eV larger than that obtained from CIS calculations.

In addition to differences in the absolute values of $H_2PO_4^-$ and $CH_3HPO_4^-$ ionization potentials, the differences in the CASPT2 and CASSCF versus CIS exciation energies lead to differences in the energetic ordering of IPs. For $H_2PO_4^-$, results from CASPT $10e10a/9e10a$ and $8e8a/7e8a$ calculations indicate that the $^2B_2$ ionization potential is smaller than that of $^2A_1$ by 0.34 and 0.38 eV, respectively. The same ordering is predicted by CASSCF $8e8a/7e8a$ and $10e10a/9e10a$ calculations. In constrast, the MP2/CIS calculations predict that the $^2A_1$ ionization potential (6.36 eV) is 0.88 eV smaller than $^2B_2$ [12, 18].

For $CH_3HPO_4^-$, differences between the CASPT2 and MP2/CIS results are similar to those occurring for $H_2PO_4^-$. CASPT2 $8e8a$/av.$7e8a$ calculations predict the a^2A' ionization potential (5.66 eV) is 0.47 eV smaller than that of b^2A'. Here, the MP2/CIS calculations predict that the b^2A' ionization potentials (6.20 eV) is 0.63 eV smaller than the a^2A' ionization potential. Like $H_2PO_4^-$, the second through fourth IPs obtained from the CASPT2 calculations are smaller than those obtained from MP2/CIS calculations. The CASPT2 $10e10a/9e10a$ value for the third IP, and the CASPT2 $10e10a/$av.$9e10a$ values for the second and fourth IPs are smaller than the MP2/CIS values by 1.28, 1.04, and 0.22 eV, respectively.

For $(CH_3)_2PO_4^-$, the ordering of the first four IPs obtained from SCF, and from $8e8a/7e8a$ CASSCF and CASPT2 calculations is the same as that predicted for $H_2PO_4^-$. Like results obtained from $H_2PO_4^-$, the CASSCF and CASPT2 ionization potentials for $(CH_3)_2PO_4^-$ are smaller than the SCF ionization potentials. The differences are between 0.29 and 0.78 eV. Also like $H_2PO_4^-$, the difference between corresponding $(CH_3)_2PO_4^-$ ionization potentials obtained from CASSCF and CASPT2 $8e8a/7e8a$ calculations is smaller than the differences between either the CASSCF or CASPT2 values, and the SCF values. The differences between IPs obtained from the CASSCF and CASPT2 calculations, with the $8e8a/7e8a$ wave functions, are in the ranges of 0.12–0.20 eV and 0.13–0.37 eV, for $H_2PO_4^-$ and $(CH_3)_2PO_4^-$, respectively.

Figures 2–6 show electron density difference maps (electron holes) associated with the five lowest IPs of $H_2PO_4^-$. The electron holes were obtained by subtracting the electron density of the radical calculated at the CASSCF $9e/10a$ level from the electron density of the anion calculated at the CASSCF $10e/10a$ level. The figures correspond to cross sections through two planes. One contains the P atom and the negatively charged O atoms (O_1 and O_2). The second plane passes through O_1 and O_2 and is perpendicular to the first plane. The electron holes associated with creation of the five lowest energy states of the $H_2PO_4^{\cdot}$ radical (a^2B_1, $^2A_2$, $^2B_2$, $^2A_1$, and b^2B_1) are similar to those predicted from earlier Koopmans' analysis of results from SCF calculations on $H_2PO_4^-$ with the 6-31G basis set, and to MP2/6-31 + G* and MP2/CIS results on $H_2PO_4^-$ and $H_2PO_4^{\cdot}$ [12].

Besides descriptions of the loss in electron density which occurs in the transitions from $H_2PO_4^-$ to $H_2PO_4^{\cdot}$, the CASSCF electron density difference maps of Figures 2–6 show regions where electron density increases. For example, in the electron distribution associated with the a^2B_1 ionization potential of $H_2PO_4^-$ (Fig. 2), the electron density increases in the region around phosphorus and in some regions where there is large contribution from $2p$-bonding and $2p$-lone-pair orbitals on the negatively charged O_1 and O_2 atoms. For the $^2A_2$ ionization potential (Fig. 3), the electron density increases in orbitals contributing to O_1PO_2 and O_1O_2 bonding interactions. Similar regions of increased electron density are observed in electron holes associated with the $^2B_2$, $^2A_1$, and b^2B_1 ionization potentials.

An earlier investigation [12] provided evidence that MP2 calculations with the 6-31 + G* basis set provide generally good predictions of the first ionization potentials of phosphorus and oxygen containing anions. With this in mind, it is likely that a combination of earlier reported MP2 results and results from CASPT2 calculations provide reasonably accurate values for the second and higher ionization potentials. Table VI lists values of the lowest energy four and five IPs obtained from

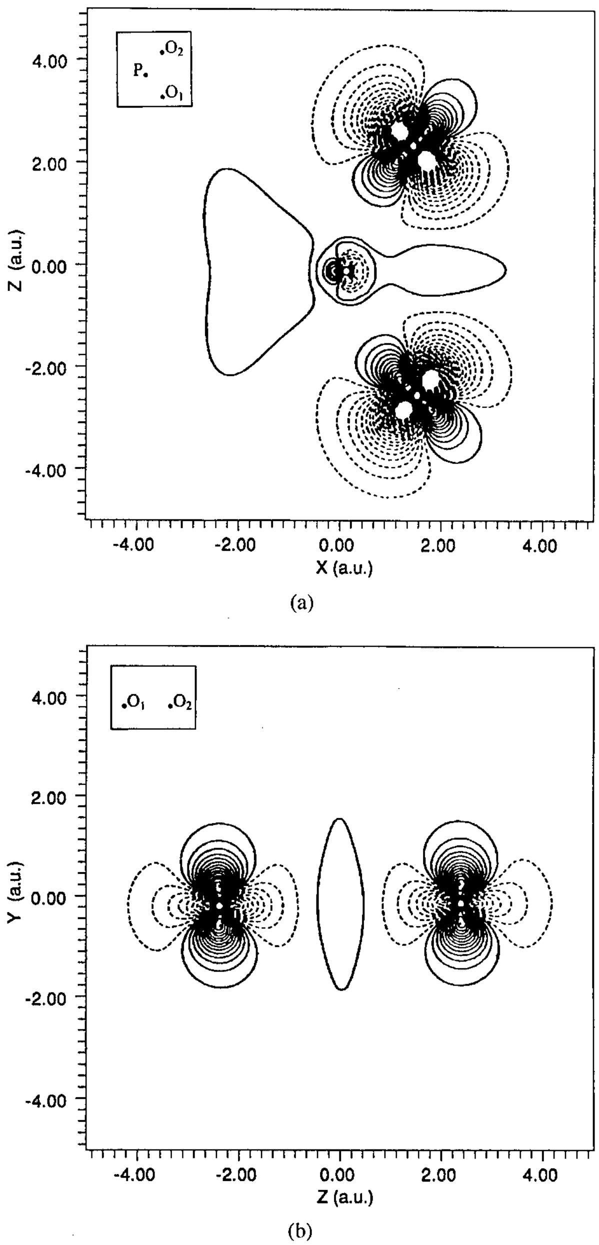


FIGURE 2. Electron hole associated with the state a^2B_1 of $H_2PO_4^{\cdot}$: (a) through the plane PO_1O_2; (b) through the plane parallel to y, 0, z and passing through O_1 and O_2. Dashed lines indicate regions in which charge density was removed on going to the radical, and solid lines indicate regions in which there is increased charge density in the radical.

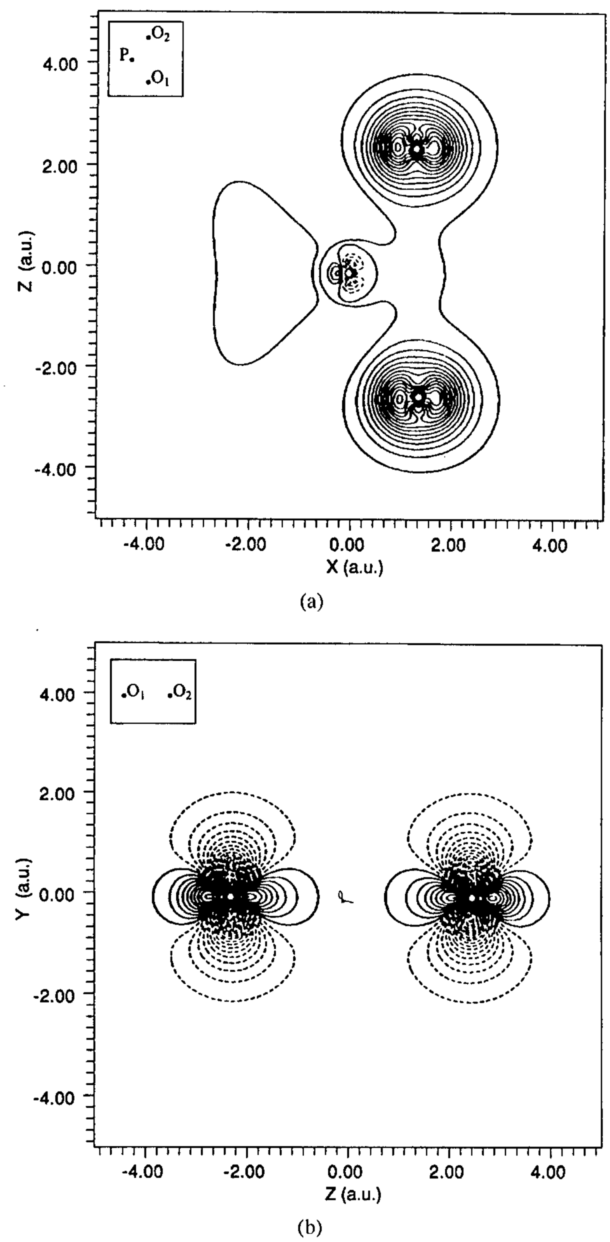


FIGURE 3. Electron hole assoicated with the state ${}^2A_2$ of $H_2PO_4^{\cdot}$: (a) through the plane PO_1O_2; (b) through the plane parallel to y, 0, z and passing through O_1 and O_2.

MP2/6-31 + G* and CASPT2 calculations for $CH_3HPO_4^-$ and $H_2PO_4^-$, respectively. Here, the first IP is taken from earlier MP2/6-31 + G* results [12, 18, 19] and the radical excitation energies are taken from the CASPT2 results. In this case, the values for the second ionization potential of $H_2PO_4^-$ (${}^2A_1$), and the corresponding b^2A' ionization potential of $CH_3HPO_4^-$ differ from the earlier reported MP2/CIS ionization potentials by no more than 0.35 eV. However, in Table VI, the values for the third, fourth, and fifth ionization potentials of $H_2PO_4^-$ (${}^2A_2$, ${}^2B_2$, and b^2A_1), and the third and forth ionization potentials of $CH_3HPO_4^-$ (b^2A'' and a^2A') are 0.49 to 1.31 eV smaller than the MP2/CIS values. For the first IPs of $H_2PO_4^-$ and

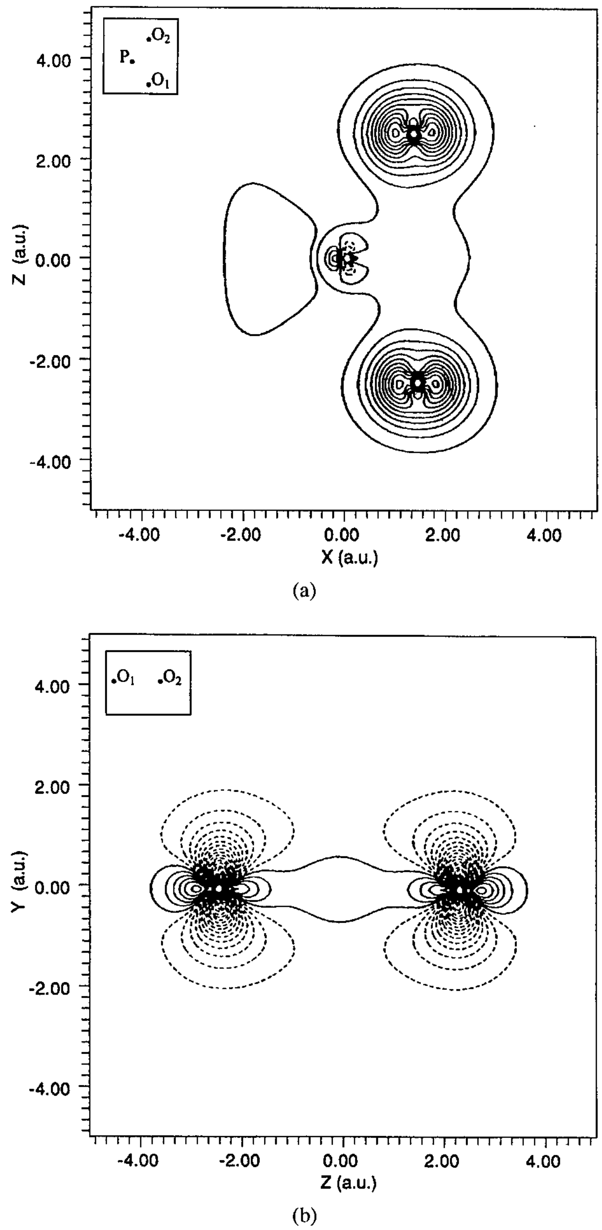


FIGURE 4. Electron hole associated with the state $^2B_2$ of $H_2PO_4^{\cdot}$: (a) through the plane PO_1O_2; (b) through the plane parallel to y, 0, z and passing through O_1 and O_2.

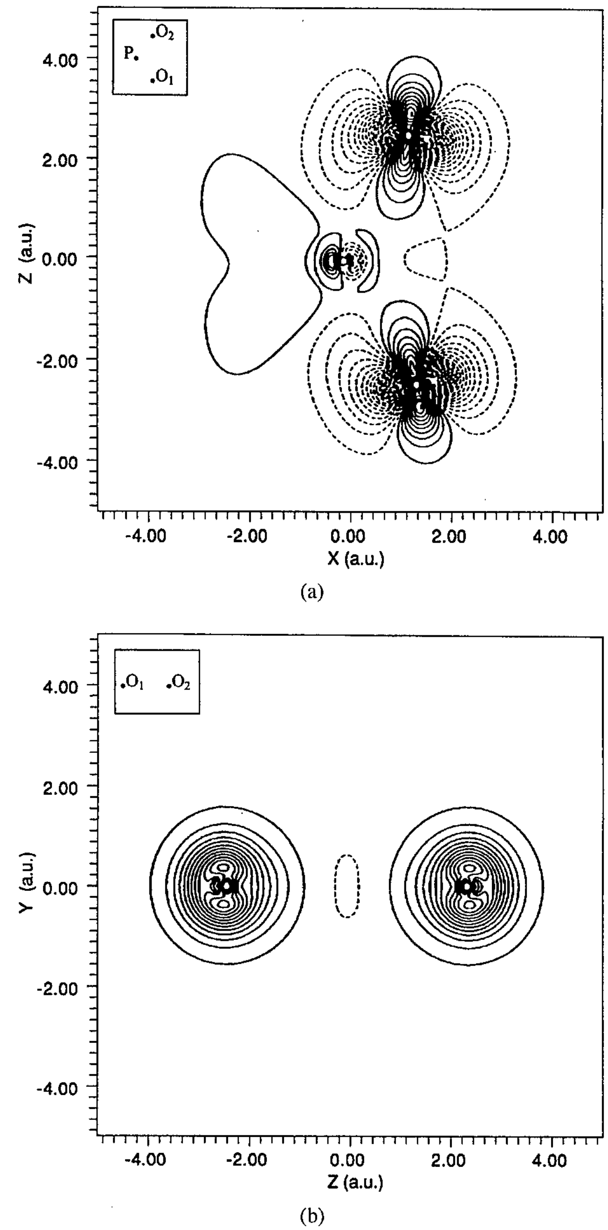
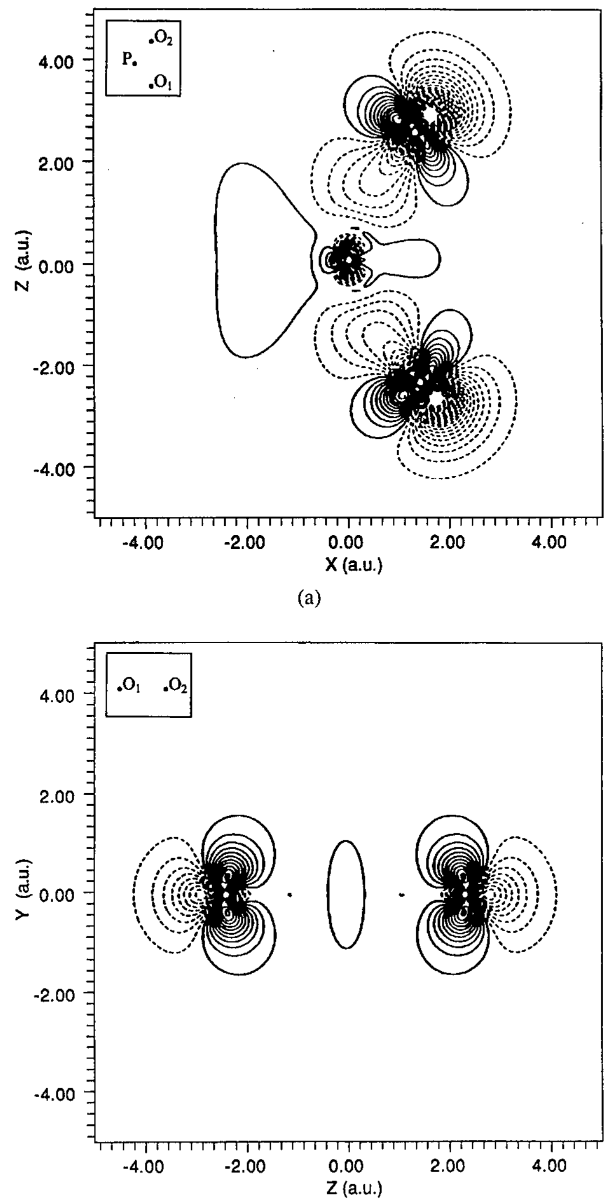


FIGURE 5. Electron hole associated with the state $^2A_1$ of $H_2PO_4^{\cdot}$: (a) through the plane PO_1O_2; (b) through the plane parallel by y, 0, z and passing through O_1 and O_2.

$CH_3HPO_4^-$, the uncertainty of ± 0.3 eV given in Table VI is based on a comparison of MP2/6-31 + G* results with experimental IPs for CH_3O^-, PO_2^- and PO_3^-. The uncertainty associated with the second, third, and fourth IPs of $H_2PO_4^-$ and $CH_3HPO_4^-$ (± 0.5 eV) reflects the uncertainty in the MP2/6-31 + G* values for the first IPs, and the variation in values of the radical excitation energies obtained from both CASSCF and CASPT2 calculations. Similarly, the larger uncertainty (± 0.8eV) assigned to the fifth IP of $H_2PO_4^-$ is based on the wider variation in CASPT2 and CASSCF results assoicated with the excitation energy for the b^2B_1 state of $H_2PO_4^{\cdot}$.

Conclusions

1. The lowest energy vertical ionization potentials of $H_2PO_4^-$ and $CH_3HPO_4^-$ predicted from CASPT2 $10e10a/9e10a$ calculations are 4.33 and 4.27 eV, respectively. For $(CH_3)_2PO_4^-$, the lowest energy IP obtained from CASPT2 $8e8a/7e8a$ calculations is 4.20 eV. The CASPT2 $10e10a/9e10a$ values of the first ionization potentials of $H_2PO_4^-$ and $CH_3HPO_4^-$ are 0.56 and 0.42 eV smaller than values, which were obtained from previously reported MP2 calculations with the 6-31 + G* basis set. An earlier comparison of MP2/6-31 + G* ionization potentials of phosphorus and oxygen containing anions [12] suggests that the MP2/6-31 + G* values for the first IPs are more accurate.
2. Significant differences in the CASSCF and CASPT2 excitation energies for $H_2PO_4^{\cdot}$ and $CH_3HPO_4^{\cdot}$ versus earlier reported energies obtained from CIS calculations with a 6-31 + G* basis set lead to a different ordering of $H_2PO_4^-$ and $CH_3HPO_4^-$ ionization potentials. Results from the MP2/CIS calculations indicate that, for $H_2PO_4^-$ ($CH_3HPO_4^-$), the $^2B_2(a^2A')$ ionization potential is greater than the $^2A_1$ (b^2A') ionization potential. Results from the CASSCF and CASPT2 calculations predict that these orderings are reversed.
3. While the CASPT2 ionization potentials of $H_2PO_4^-$ are smaller than earlier reported IPs obtained from MP2/6-31 + G* and MP2/CIS calculations, and the ordering of IPs is different, the descriptions, which the current and the earlier results provide of electron holes with corresponding symmetries are qualitatively similar to one another. These descriptions are also similar to those obtained by applying Koopmans' theorem to results from split-valence basis set SCF calculations on the ground-state anions.
4. Currently, the most accurate values of the four and five lowest energy IPs of $CH_3HPO_4^-$ and $H_2PO_4^-$, respectively, are believed to arise from a combination of results from MP2/6-31 + G* calculations of the first IPs with results from CASPT2 calculations of $CH_3HPO_4^{\cdot}$ and $H_2PO_4^{\cdot}$ excitation energies.



(b)

FIGURE 6. Electron hole associated with the state b^2B_1 of $H_2PO_4^{\cdot}$: (a) through the plane PO_1O_2; (b) through the plane parallel to y, 0, z and passing through O_1 and O_2.

ACKNOWLEDGMENTS

Support of this work by the American Cancer Society (Grant #CN-37E) is gratefully acknowledged. Computer access time has been provided by the CNRS, the Computer Center of The Univer-

TABLE VI
Ionization potentials of $H_2PO_4^-$ and $CH_3HPO_4^-$ obtained by combining MP2 / 6-31 + G* and CASPT2 results.

	Ionization potentials[a]				
	IP_1	IP_2	IP_3	IP_4	IP_5
$H_2PO_4^{-}$ [b]	4.89 ± 0.3	5.42 ± 0.5	6.30 ± 0.5	6.64 ± 0.5	7.41 ± 0.8
$CH_3HPO_4^{-}$ [c]	4.69 ± 0.3	5.43 ± 0.5	6.08 ± 0.5	6.55 ± 0.5	

[a] In electron volts. MP2 / 6-31 + G* results obtained from Refs. [12], [18], and [19]. The estimated uncertainty in the reported values is discussed in the text.
[b] The first five IPs associated with formation of the a^2B_1, $^2A_2$, $^2B_2$, $^2A_1$, and b^2B_1 radical states, respectively. Obtained from CASPT2 10*e*10*a* / 9*e*10*a* results.
[c] Ionization potentials associated with formation of the a^2A'', b^2A'', and a^2A' and b^2A' radical states. Ionization potentials for the b^2A'' and b^2A' states obtained from CASPT2 10*e*10*a* / av.9*e*10*a* results. Ionization potentials for the a^2A' state obtained from CASPT2 8*e*8*a* / av.7*e*8*a* results.

sity of Illinois at Chicago, the Cornell Theory Center, and the National Center for Supercomputing Applications, at the University of Illinois at Urbana-Champaign.

References

1. P. R. LeBreton, X. Yang, S. Urano, S. Fetzer, M. Yu, N. J. Leonard, and S. Kumar, J. Am. Chem. Soc. **112**, 2138 (1990).
2. S. Urano, X. Yang, and P. R. LeBreton, J. Mol. Struc. **214**, 315 (1989).
3. A. O. Colson, B. Besler, and M. D. Sevilla, J. Phys. Chem. **97**, 13852 (1993).
4. A. O. Colson, B. Besler, and M. D. Sevilla, J. Phys. Chem. **99**, 1060 (1995).
5. K. Szczepaniak, M. Szczesniak, W. Szajda, W. B. Person, and J. Leszczynski, Can. J. Chem. **69**, 1705 (1991).
6. M. Szczesniak, J. Leszczynski, and W. B. Person, J. Am. Chem. Soc. **114**, 2731 (1992).
7. W. B. Person and K. Szczepaniak, in *Vibrational Spectra and Structure*, Vol. 20, J. R. Durig, Ed. (Elsevier Science, Amsterdam, 1993), pp. 239–325.
8. P. E. Young, I. H. Hiller, and I. R. Gould, J. Chem. Soc. Perkin Trans. 2, 1717 (1994).
9. M. Yu, H. S. Kim, and P. R. LeBreton, Biochem. Biophys. Res. Commun. **184**, 16 (1992).
10. A. O. Colson, B. Besler, D. M. Close, and M. D. Sevilla, J. Phys. Chem. **96**, 661 (1992).
11. A. O. Colson, B. Besler, and M. D. Sevilla, J. Phys. Chem. **97**, 8092 (1993).
12. H. S. Kim, M. Yu, Q. Jiang, and P. R. LeBreton, J. Am. Chem. Soc. **115**, 6169 (1993).
13. K. Tasaki, X. Yang, S. Urano, S. Fetzer, and P. R. LeBreton, J. Am. Chem. Soc. **112**, 538 (1990).
14. G. P. Ford and J. D. Scribner, Chem. Res. Toxicol, **3**, 219 (1990).
15. G. P. Ford and B. Z. Wang, Carcinogenesis **14**, 1465 (1993).
16. L. Pardo, R. Osman, H. Weinstein, and J. R. Rabinowitz, J. Am. Chem. Soc. **115**, 8263 (1993).
17. D. P. Vercauteren and E. Clementi, Int. J. Quant. Chem.: Quant. Biol. Symp. **10**, 11 (1983).
18. H. S. Kim and P. R. LeBreton, Proc. Natl. Acad. Sci. USA **91**, 3725 (1994).
19. N. S. Kim and P. R. LeBreton, J. Am. Chem. Soc. **118**, 3694 (1996).
20. N. S. Kim, Q. Jiang, and P. R. LeBreton, Int. J. Quant. Chem.: Quant. Biol. Symp. **23**, 11 (1996).
21. N. S. Kim and P. R. LeBreton, Biospectroscopy, **3**, 1 (1997).
22. T. Koopmans, Physica **1**, 104 (1934).
23. D. Dougherty, K. Wittel, J. Meeks, and S. P. McGlynn, J. Am. Chem. Soc. **98**, 3815 (1976).
24. L. Asbrink, C. Fridh, and E. Lindholm, Tetrahedron Lett. **52**, 4627 (1977).
25. D. Perahia, A. Pullman, and H. Berthod, Theoret. Chim. Acta **40**, 47 (1975).
26. H. Berthod and A. Pullman, Chem. Phys. Lett. **32**, 233 (1975).
27. M. D. Newton, J. Am. Chem. Soc. **95**, 256 (1973).
28. W. J. Hehre, L. Radom, P. V. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1986) pp. 38–40, 79, 82, 86.
29. J. Foresman, M. Head-Gordon, J. A. Pople, and M. J. Frisch, J. Phys. Chem. **96**, 135 (1992).
30. P. E. M. Siegbahn, J. Almlöf, A. Heiberg, and B. O. Roos, J. Chem. Phys. **74**, 2384 (1981).
31. K. Andersson, P. A. Malmquist, B. O. Roos, A. J. Sadlej, and K. J. Wolinski, J. Phys. Chem. **94**, 5483 (1990).
32. K. Andersson, P. A. Malmquist, and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).
33. B. T. Cruse, E. Egert, O. Kennard, G. B. Sala, S. A. Salisbury, and M. A. Visamitra, Biochemistry **22**, 1833 (1983).
34. C. Lim and M. Karplus, J. Am. Chem. Soc. **112**, 5872 (1990).
35. A. Dejaegere, C. Lim, and M. Karplus, J. Am. Chem. Soc. **113**, 4353 (1991).
36. B. Ma, Y. Xie, M. Shen, and H. F. Schaefer, J. Am. Chem. Soc. **115**, 1943 (1993).

37. A. Veillard, Theoret. Chim. Acta **12**, 405 (1968).

38. B. O. Roos and P. Siegbahn, Theoret. Chim. Acta **17**, 199 (1970).

39. D. M. Hayes, G. L. Kenyon, and P. A. Kollman, J. Am. Chem. Soc. **100**, 4331 (1978).

40. H. Wallmeier and W. Kutzelnigg, J. Am. Chem. Soc. **101**, 2804 (1979).

41. R. Höller and H. Lischka, J. Am. Chem. Soc. **102**, 4632 (1980).

42. S. Huzinaga, *Approximate Atomic Functions*, Technical Report (University of Alberta, Alberta, 1971).

43. S. Huzinaga, J. Chem. Phy. **42**, 1293 (1965).

44. P. A. Malmqvist, B. O. Roos, M. P. Fülscher, and A. P. Rendell, Chem. Phys. **162**, 359 (1992).

Pharmacophoric Pattern in Valpromide Derivatives

SILVINA TASSO,[1] LUIS BRUNO-BLANCH,[1] GUILLERMINA L. ESTIÚ[2]

[1] *Química Medicinal, División Farmacia, Facultad de Ciencias Exactas, UNLP, Casilla de Correo 243-1900-La Plata, Argentina*

[2] *Cequinor, Facultad de Ciencias Exactas, UNLP, Casilla de Correo 962-1900-La Plata, Argentina*

Received 5 March 1997; accepted 16 May 1997

ABSTRACT: A conformational and electronic semiempirical quantum-chemical study of several *N*-substituted valpromides is presented, followed by a similarity analysis that takes into account the flexibility of the molecules. Rigid analogs are included in the comparison in order to help identify the anticonvulsant active conformations. On the basis of a superposition analysis, which includes both active and nonactive structures, and uses the global minimum-energy conformation of phenytoin as a template, the pharmacophoric pattern of *N*-substituted valpromides is defined. It is related to the antiperiplanar orientation of the amide function relative to the hydrocarbon chain.
© 1997 John Wiley & Sone, Inc. Int J Quant Chem **65**: 1107–1114, 1997

Key words: Valpromide derivatives; anticonvulsant activity; similarity analysis; pharmacophore

Introduction

Since the introduction of carbamazepine and valproate into clinical practice during the 1970s, the clinical science of epilepsy has progressed considerably. Nowadays, epileptic seizures are most commonly treated with one of four antiepileptic drugs (AED): carbamazepine (cbz, 31%), phenytoin (phen, 31%), valproate (vpa, 26%), and phenobarbital (pb, 18%) [1–3]. New promising compounds, such as felbamate [4, 5], gabapentin, remacemide, and tiagabine [6], among others, have been also recently marketed.

The optimal goal for epilepsy treatment is the complete control of seizure with no adverse effects. However, the achievement of a balance between seizure control and adverse effects, which include neurotoxicity, hepatotoxicity, and even teratogenicity [1, 2, 6], still remains a challenge for the most expert scientist. The success in developing more effective AEDs in the future will depend, in large part, on our ability to broaden our under-

Correspondence to: G. L. Estiú.

Contract grant sponsors: CONICET, Universidad Nacional de La Plata, Cooperativa Farmacеútica de Quilmes (COFARQUIL), Laboratorios Bagó, and Colegio de Farmacéuticos de la Provincia de Buenos Aires.

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061107-08

standing of the molecular mechanisms of epileptogenesis. Molecular similarity studies, either based on the superimposition of nuclei, the comparison of electron density matrices, or on any other principle [7], have been extremely useful for this purpose, as they allow one to identify the structural and electronic requirements associated with a particular activity. The overall picture is, however, complicated by the different pharmacokinetic (metabolic) paths that similar structures might undergo in the biological media. A typical example of this complicated picture is given by vpa derivatives. Vpa is recognized as a first-line AED, because of its broad spectrum of activity and its lower neurotoxicity [1, 8]. The primary amide of vpa, valpromide (vpd) has been found to be 2–5 times more potent than vpa in mice [9, 10]. Pharmacokinetic studies seem to indicate, on the other hand, that it is biotransformed to the acid, to the extent of 30–40% in dogs [11, 12], and 80% in humans [13].

In spite of the complication related to their different metabolic behavior in the different species, the amides represent an interesting alternative because they are more active as anticonvulsants, less teratogenic and probably less hepatotoxic than the respective acids [14]. Several analogs of vpd have been studied in their pharmacokinetics and for some of them (e.g., valnoctamide) the anticonvulsant activity has been determined. From the results derived from those studies, it has been inferred that the amides that do not undergo biotransformation to the acid are the most active ones [14, 15]. This statement implies that the pharmacokinetics is the limiting factor for the anticonvulsant activity.

In the systematic analysis that has led to that conclusion, only two *N*-substituted derivatives (*N*-methyl tetramethylcyclopropane carboxamide and tetramethyl cyclopropylcarbonylglycinamide) have been included [15]. However, to our knowledge, neither the pharmacokinetics nor the anticonvulsant activity of *N*-substituted valpromides have been studied in a systematic way, oriented to discern the influence of *N*-substitution on the AE activity. We have focused on this group of molecules and, in addition to the synthesis and biological testing of the AE activity, we have performed a similarity analysis, based on the superposition of conformers followed by the comparison of their charges, calculated by means of quantum-chemical methodologies. This analysis was oriented to assess the structural and electronic requirements associated with the activity.

This article mainly concerns the details of the similarity analysis for a series of *N*-substituted valpromides, which have been, as mentioned before, synthesized and biologically tested in our lab. The comparative study, which includes both active and inactive structures, as well as rigid analogs selected to confirm the inferences derived from the conformational analysis, can be defined as a structure activity relationship (SAR) analysis and has allowed us to identify the pharmacophoric pattern for the AE active *N*-substituted valpromides.

Methodologies: Calculation Procedure

Six *N*-substituted valpromides (*N*-butyl-vpd, *buvpd*; *N,N*-diethyl-vpd, *dievpd*; *N*-cyclohexyl-vpd, *chvpd*; *N*-isopropyl-vpd-, *ipvpd*; *N*-(4-carboxyphenyl)-vpd, *cpvpd*; and *N*-morpholin-vpa, *mvpd*) have been synthesized and biologically tested for their AE activity. In order to gain insight in the microscopic origin of the manifested activity, a molecular similarity analysis was performed after a thorough quantum-chemical conformational study. Similarity was based on the comparison of both geometric and electronic descriptors, defined, respectively, by the position of the nuclei and the charges on the atomic centers, taking into account that the receptor site perceives charge distributions of the approaching molecule [16]. Rigid analogs, rigidified by cyclation, have been included in the comparison, in order to help discern the conformation of the flexible molecules.

Details of the synthesis and pharmacological studies are given in Ref. [17]. In a general scheme, the synthesis can be described as an SN1 reaction of valproyl chloride with *N*-butyl, *N,N*-diethyl, *N*-cyclohexyl, *N*-isopropyl, *N*-(4-carboxyphenyl)-amide and morpholine, respectively.

The pharmacologic tests were performed according to standard procedures provided by the Antiepileptic Drug Development Program of the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) [18, 19]. Maximal electroshock seizure (MES) and pentylentetrazol seizure threshold (PTS) tests were employed to determine the anticonvulsant activity. Rotorod test was used to determine the acute toxicity. The AE activity was expressed as ED50 (the dose that is effective in 50% of the animals

tested), and estimated, with their 95% confidence limits, by probit analysis [20].

Because the size of the molecules is not compatible with good quality ab initio calculations, an AM1 model Hamiltonian [21] (MOPAC 7.0 package [22]) was chosen for the conformational search. For each derivative, the structures associated with the initial guesses for a gradient-driven full geometry optimization were generated by means of modifications of the torsional angles τ_5 and τ_6 (Fig. 1), and of those defined in the hydrocarbon chain ($\tau_1-\tau_4$). These and the other geometry parameters were completely relaxed during the optimizations.

The conformational search has been performed as follows:

1. The τ_5 value was modified in 90° steps from 0° to 180° (270° is equivalent to 90° for this set of molecules). Intermediate values were not considered because all the optimizations starting from the above-mentioned ones converged to values close to either $\tau_5 = 0°$ or $\tau_5 = 180°$.

2. For each of the τ_5 values, τ_6 has been varied in 90° steps. In a similar fashion to that described for τ_5, two minima were found, associated, respectively, with the orientation of the less voluminous substituent (a hydrogen atom in the case of single *N*-substitution) toward the hydrocarbon chain or opposite to it. The first one is the most stable because it minimizes steric repulsion.

3. As will be further discussed, it is well known that the "all-trans" conformation is the most stable for the hydrocarbon chain. A thorough discussion of this subject can be found in Ref. [23]. This conformation has been confirmed, however, for the different derivatives, by means of distortions of the $\tau_1-\tau_4$ angles in 60° and 90° from their starting 180°, followed by full optimization of the resulting structure.

4. Several minima were defined, after the geometry optimization procedure, which were associated with the local conformation of the *N*-substituents. The largest number of minima has been found for the most flexible molecule, *N*-butil-valpromide: three minima related to the coordinates of the substitutent have been found for each pair of values of τ_5 and τ_6, with an energy difference among them lower than 3 kcal/mol. The one of lowest energy has been always considered in further steps of the study.

We have chosen AM1 as the calculation procedure, on the basis of the comparison of the results of semiempirical AM1, PM3, MNDO [22], and ZINDO/S [24, 25] gas-phase calculations, the latter at the configuration interaction singles (CIS) and singles and doubles (CISD) levels, with those of ab initio calculations of different quality (G94 [26]), for a smaller molecule, acetamide [27], which is representative of the set. In the semiempirical calculations the geometry of the molecule has been optimized within each methodology. Only for the ZINDO/S calculations the AM1 optimized structure was used. The ab initio computational methodology uses Gaussian 94 [26], at the 6-31G(d), 6-31G(d, p), 6-31 + G(d), 6-31 + G(d, p) and 6-311 + G(d, p) levels of theory. The influence of electron correlation was analyzed at the MP2 and CISD levels, for the geometry optimized at the 6-311 + G(d, p) HF level. In each case, orbital-based descriptions (Mulliken population analysis, MPA [28], and natural population analysis, NPA [29]) and charges from electrostatic potentials (CHelpG [30]) derived from ab initio calculations, have been compared with the MPA that follows each semiempirical approach. Being aware of the fact that the local charges are not quantum mechanical observables, the calculated dipole moments were compared with the experimental value (3.76 D [31]), as a way of testing the accuracy of the calculations. From the comparison of the equilibrium geometries, electronic properties, and even the description of the resonant effect in acetamide [27, 32], we conclude that the AM1 results are the closest to ab initio, giving also a calculated value for the dipole moment very close to experiment.

As will be further explained in the next section, the value of τ_5 differentiates two well-defined conformations that are also characterized by a dif-

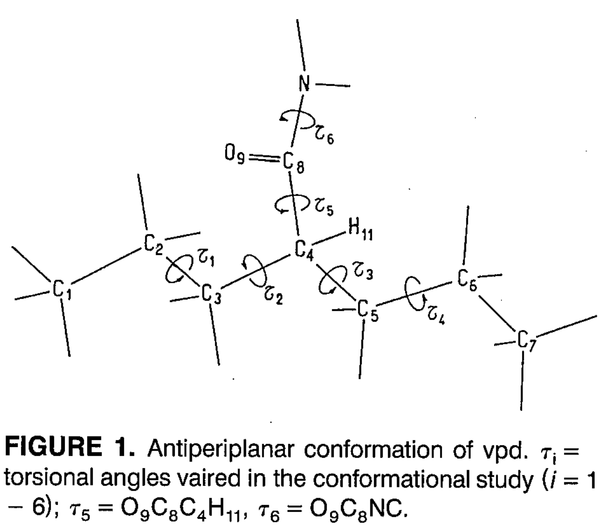


FIGURE 1. Antiperiplanar conformation of vpd. τ_i = torsional angles vaired in the conformational study (i = 1 − 6); $\tau_5 = O_9C_8C_4H_{11}$, $\tau_6 = O_9C_8NC$.

ferent pharmacological response. AM1 calculations have been also applied to calculate the energy barriers involved in the mutual interconversion between those conformations for each derivative. To this end, the torsional angle that defines the reaction coordinate to evolve among the different conformations (τ_5) has been varied in 10° steps. For each step, this coordinate was kept frozen while the others were fully optimized.

The selection of AM1 as the semiempirical procedure for the study of systems of biological interest, including searches for transition states, has been also validated against ab initio calculations by R. Cachau [33].

Electronic descriptors have been derived from a Mulliken population analysis [28] performed at the AM1 level. In spite of the lack of precision of this analysis for absolute calculations, their results are widely accepted, in this field, for the study of the trends in their variation, on well-defined atomic centers, that follow structural modifications performed to a parent structure [23, 24, 35].

TABLE I
Anticonvulsant activity, expressed as DE_{50}, of the compounds analyzed.[a]

	ED_{50} PTZ test Dose (μmol/kg)	ED_{50} MES test Dose (μmol/kg)
vpa	1261 (1155–13771)	751 (526–1074)
vpd	385[b]	392[b]
phen	NE[c]	24[c]
cpvpd	1847 (1491–2288)	1159 (1002–1340)
buvpd	NE (a 375 μmol/kg)	91 (55–152)
chvpd	NE (a 200 μmol/kg)	66 (36–121)
mvpd	NE (a 1550 μmol/kg)	NE (a 1500 μmol/kg)
dievpd	NE (a 375 μmol/kg)	128 (93–176)
ipvpd	NE (a 200 μmol/kg)	142 (114–177)

[a] DE_{50}: 50% effective dose, NE: not effective. vpa, valproic acid; vpd, valpramide; phen, phenytoin; cpvpd: *N*-(4-carboxyphenyl); buvpd: *N*-buty; chvpd: *N*-cycloheyl; mvpd: *N*-morpholin; dievpd: *N*,*N*-diethyl ipvpd: *N*-isopropylvalparamide.
[b] S. Hadad, T. Vree, E. Kleijn, and M. Bialer, J. Pharm. Sci. **81**(10) (1992) 1047.
[c] E. Shek, T. Murakami, C. Nath, E. Pop, and N. Bodor, J. Pharm. Sci. **78** (10) (1989) 837.

Results and Discussion

From the comparison of the AE activity (Table I), it becomes noticeable that, whereas almost all the substituted amides (except mvpd) are active against MES, they are not (except cpvpd) active against PTS. The different response against both types of convulsions defines a difference between the derivatives vs. vpa and vpd (Table I) and is indicative of a different reaction mechanism. Within the MES model, buvpd, dievpd, chvpd, ipvpd, and cpvpd manifest AE activity, which is higher, in general, then vpd. However, one of them, the morpholin derivative, does not show protection at all against convulsion. This difference may contain the clue to understand the requirements associated with the AE activity. We have devoted our research to find the origin of this difference.

In agreement with previous calculations by us [23] and other authors [36], we have found the larger stability for the linear, all-trans conformation in the valproyl moiety (Fig. 1). Further confirmation is obtained from X-ray diffraction studies of one of the derivatives of the series (cpvpd) [37].

In the carboxamide moiety, on the other hand, two minima were found after the geometry optimization procedure, which are related to values of τ_5 close to 0° and 180°, respectively. This fact allows one to distinguish between an "eclipsed" (synperiplanar) and an "opposite" (antiperiplanar) O_9 and H_{11} conformations (Fig. 1). According to the results of the AM1 calculations (Table II), the synperiplanar conformation (τ_5 close to 0°) is preferred by several active molecules (vpd, cpvpd, buvpd, chvpd, ipvpd), whereas the nonactive, morpholin derivative is more stable in the opposite one. Far from being a conformational requirement, the energy difference between both orientations, lower in general than 1 kcal/mol (increasing to about 2 kcal/mol for dievpd and ipvpd), shows that both conformers can coexist in equilibrium, provided that both are formed as a result of their synthesis. The value of τ_6 also define two different conformations. The most stable one is always related to the orientation of the less voluminous *N*-substitutent toward the hydrocarbon chain, for both the synperiplanar and the antiperiplanar con-

TABLE II
Relevant conformational data (dihedral angles, τ_5, τ_6) of the most stable geometries of the structure analzyed.

	Dihedral angles		ΔE	b.h
	τ_6	τ_5	(kcal)	(kcal)
vpa	0	−179.3	0.7	2.1
vpd	−0.22	0.46	0.8	2.2
phen	179.90	−176.28	–	–
cpvpd	5.87	−0.95	0.7	2.3
buvpd	−7.93	2.03	0.9	2.1
chvpd	−0.35	−2.02	0.8	2.1
mvpd	−1.11	169.23	1.3	19.0
dievpd	6.93	167.58	2.3	6.56
ipvpd	−0.10	−2.60	1.75	2.33

[a] Energy differences (ΔE) between synperiplanar and antyperiplanar conformations, and energy required (b.h) for their mutual interconversion. $\tau_6 = O_9C_8NC$, $\tau_5 = O_9O_8C_4H_{11}$.

formations. However, no distinguishable pharmacologic activity is associated with them.

Because we have found the active–nonactive condition related to the value of the τ_5 angle, we have centered our study in the energetics associated with the mutual interconversion of the conformers defined by it. The calculated energy barriers (Table II) are very small for the active derivatives ($\sim$ 2 kcal/mol), but the barrier increases to more than 10 times for the nonactive morpholin, due to steric repulsion between the voluminous morpholin substitutent and the hydrocarbon chain in the intermediate configuration of the reaction path, defined by $\tau_5 = 0°$. It is noticeable that both the characteristics of the most stable conformation, and the energy difference between the more stable ones, seem to be determined by a steric factor, associated with the size of the *N*-substitution. In the set of compounds that has been considered, single *N*-substitution has never led to nonactive structures. Disubstitution increases the barrier in both dievpd and mvpd, but the calculated barrier of the former is not high enough to consider it as unable to accommodate to the requirements of the receptor. In agreement with this, the biological test (Table I) indicates that dievpd is AE active.

The active molecules are characterized, thence, by soft degrees of rotational freedom and can accommodate themselves to the conformational requirements defined by the receptor site at a very low energy cost. Rotation around the C_4–C_8 bond is impeded, on the other hand, for the nonactive morpholin derivative, which will remain in the conformation that results from its synthesis. This conformation, once it is known, defines the structural characteristics of the nonactive structures.

At this point we can say nothing about the active conformation of the valpromide derivatives, because of the flexibility of the active structures, on one side, and the lack of knowledge of the conformation that results from the synthesis, on the other. The flexibility of the active molecules does not allow us to proceed to a superposition analysis based on the comparison of their structures, as all of them can be superimposed at a very low expense of energy. In order to assess the conformational characteristics of the active compounds, we have considered rigid analogs, rigidified by cyclation, whose conformation and AE activity are perfectly known [38]. To this end, we have included phenytoin (Fig. 2) in the series and used it as our template for the similarity analysis. It has a rigid antiperiplanar conformation, defined by the O_9 and N atoms. The structure is rigidified in both the hydrocarbon chain and in the amide group, and is active against the MES test (Table I). It is nonactive against PTS, a fact that can be understood as indicative of a similar AE mechanism as the *N*-substituted valpromides analyzed in this series, but cpvpd.

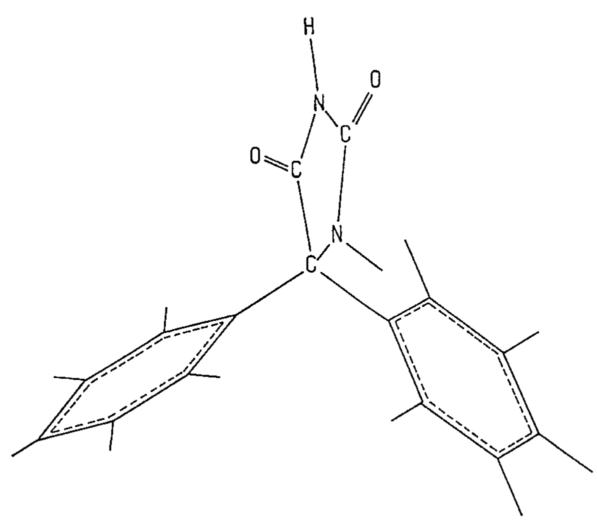


FIGURE 2. Most stable conformation of phenytoin, a rigid analog, as a result from AM1 calculations.

The structure of phenytoin partially overlaps with the antiperiplanar conformation of the substituted valpromides (Fig. 3), but not with the synperiplanar. Moreover, the charges on the atomic centers are almost the same for the overlapping portion. On the basis of this analysis, we can associate the AE activity (protection against MES) of the molecules under study with the antiperiplanar conformation. Thence, regardless of the conformation obtained by synthesis, the active structures are flexible and can adopt this active conformation without energy penalty. For the morpholin derivative, on the other hand, the synperiplanar conformation is very likely to be favored in the synthesis, because it avoids steric repulsion between the H atom of the valproyl moiety and the morpholin group. Its lack of flexibility disables it to adopt the opposite configuration, and, in this way, its lack of activity can be understood.

We want to remark that the active conformations derived from the similarity analysis based on the comparison with phenytoin (antiperiplanar) are not the most stable ones that result from the conformational search, based on AM1 calculations (Table II). However, because of the flexibility of the molecules, the conformational analysis is not sufficient per se to establish the structural requirements of the active structures. We have based our conclusions on their comparison with phenytoin.

The superposition of the structure of phenytoin with the eclipsed and opposite conformations of the substituted valpromides (Fig. 3) allows one to define a pharmacophoric pattern for the latter, which is mainly associated with the opposite orientation of the carbonylic group relative to H_{11} (antiperiplanar orientation, Fig. 4). The nonzero AE activity of the *N,N*-diethyl valpromide demonstrates that the requirement of having a H atom bonded to the aminic nitrogen is not included in the definition of the pharmacophore. According to the superposition analysis, the straight conformation of the hydrocarbon chain does not appear either as a requisite for the AE activity. The definition of the pharmacophore includes the carbon atoms of the aliphatic chain in α and β positions relative to the amide functionality. Although it does not include the seven carbon atoms of the valproil moiety, it should be kept in mind that, within a series of monocarboxilyc acids, vpa has the optimal chemical structure with regards to margins between its anticonvulsant activity and its sedative or hypnotic effects [39, 10]. Consequently, the importance of the size of the hydrocarbon chain cannot be disregarded.

The local charges on the atoms of the valproil moiety are not significantly modified by substitution. As the charges remain the same along the series, including phenytoin, structural similarity in the definition of the pharmacophore implies electronic-structural similarity. This is in agreement

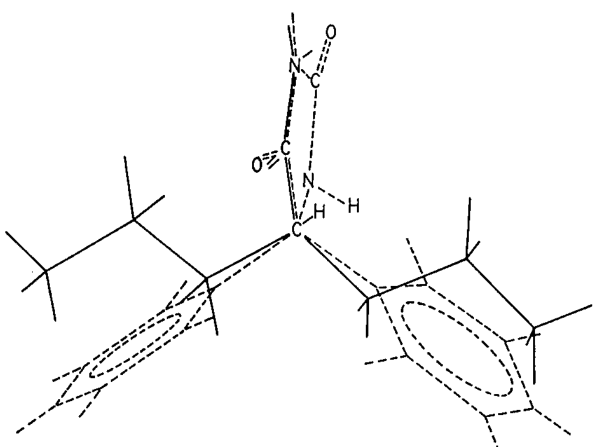


FIGURE 3. Superposition of the structure phenytoin with the opposite conformation (carbonyl group relative H_{11}) of the substituted valpromides, exemplified by vpd. A stick model has been chosen because it is the only one that allows to visualize the overlapping. Thick line: vpd; broken line: phenytoin.

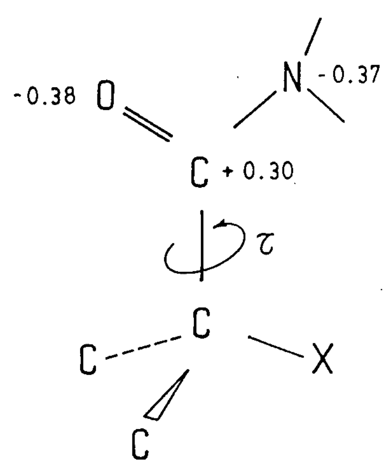


FIGURE 4. Molecular portion that define the structural requirement associated with the AE activity (pharmacophore). X refers to H for the set of substituted valpromides and to N for phenytoin. The local charges on the most relevant atoms are shown.

with the premise that the receptor site perceives electronic distributions approaching to it.

According to the previous description, our conclusions have been mainly derived from the comparison of the syn- and antiperiplanar conformers of the derivatives of the series with phenytoin (rigid analog), attending both the nuclear coordinates and the local charges on the atomic centers. The superimpossible portion of the active derivatives define the pharmacophore.

Having identified the pharmacophore associated with the AE activity of the *N*-substituted valpromides, the next step is oriented to find a structural or electronic quantifier that, correlating with the AE potency, would allow us to design new derivatives with a predetermined activity. This information has to be derived from an QSAR analysis. Because more data (derivatives) are necessary for this analysis to become statistically significant, research in our lab is presently conducted in the synthesis of new derivatives, which are being designed on the basis of the knowledge of the pharmacophoric pattern.

Concluding Remarks

In this article we have presented a conformational and electronic study of several *N*-substituted valpromides, followed by a similarity analysis that, taking into account the flexibility of the molecules, has considered a rigid analog (phenytoin) as template. Phenytoin has been included in the analysis, and both the position of the nuclei and the local charges on the atomic centers have been compared with those of the active and inactive structures. This procedure has allowed us to define a pharmacophore, related to the antiperiplanar orientation of the amide function relative to the hydrocarbon chain, and defined by the structural portion shown in Figure 4.

On the basis of the knowledge of the pharmacophoric pattern, new derivatives are being synthesized in our lab in order to be able to perform a statistically significant QSAR analysis.

ACKNOWLEDGMENT

G. L. Estiú is a member of the Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina (CONICET), and L. E. Bruno-Blanch of the Facultad de Ciencias Exactas, Universidad Nacional de La Plata. S.M.T. acknowledges Laboratorios Bagó (Argentina) for a Research Fellowship. This work was supported in part through grants from CONICET, Universidad Nacional de La Plata, Cooperativa Farmaceútica de Quilmes (COFARQUIL), Laboratorios Bagó, and Colegio de Farmacéuticos de la Provincia de Buenos Aires, Argentina. We gratefully acknowledge the use of the computer facilities of the Cátedra de Química Cuántica, Facultad de Química, Universidad de la República Oriental del Uruguay.

References

1. D. Chadwick, Epilepsia **35** (Suppl. 4), S3 (1994).
2. M. Bialer, A. Haj-Yehia, K. Badir, and S. Hadad, Pharm. World Sci. **16**(1), 2 (1994).
3. (a) E. A. Swinyard, J. H. Woodhead, H. Steve White, and M. R. Franklin, in *Antiepileptic Drugs*, R. Levy, R. Mattson, B. Meldrum, J. K. Penry and F. E. Dreyfuss, Eds. (Raven Press, New York, 1989), Chapter 5. (b) R. H. Levy and D. D. Shen in *Antiepileptic Drugs*, R. Levy, R. Mattson, B. Meldrum, J. K. Penry, and F. E. Dreifuss, Eds., 3rd ed. (Raven Press, New York, 1989), Chapter 41.
4. W. E. Dodson, Epilepsia **34** (Suppl. 7), S18 (1993).
5. D. Schmidt, Epilepsia **34** (Suppl. 7), S30 (1993).
6. J. M. Pellock, Epilepsia, **35** (Suppl. 4), S11 (1994).
7. R. Carbo Dorca, and P. J. Mezey, Eds., *Advances in Molecular Similarity, Vol. 1*, (JAI Press, Greenwich, 1996).
8. A. Richens and E. Penicca, in *Textbook of Epilepsy*, J. Lidlaw, A. Richens, and D. Chadwick, Eds. (Churchill Livingstone, Edinburg, 1992), p. 177.
9. A. Haj-Yehia and M. Bialer, J. Pharm. Sci. **79**(8), 719 (1990).
10. W. Loscher and H. Nau, Neuropharmacology **24**, 427 (1985).
11. M. Bialer and A. Rubinstein, J. Pharm. Pharmacol. **35**, 607 (1983).
12. M. Bialer and A. Rubinstein, Biopharm. Drug Dispos. **5**, 177 (1983).
13. M. Bialer, A. Rubinstein, J. Dubrovsky, Y. Raz, and O. Abramsky, Int. J. Pharm. **23**, 25 (1985).
14. A. Haj-Yehia, S. Hadad, and M. Bialer, Pharm. Res. **9**(8), 1058 (1992).
15. M. Bialer, S. Hadad, B. Kadry, A. Abdul-Hai, A. Haj-Yehia, J. Sterling, Y. Herzing, and B. Yagen, Pharm. Res **13**(2), 284 (1996).
16. L. M. Balbes, S. W. Mascarella, and D. B. Boyd, in *Reviews in Computational Chemistry*, Vol. 4, K. B. Lipkowitz and D. B. Boyd, Eds. New York, (VCH Publishers, 1994).
17. S. Tasso, L. Bruno-Blanch, and G. Estiú, in preparation.
18. R. L. Krall, J. K. Penry, B. G., White, H. J. Kuperferberg, and E. A. Swinyard, Epilepsia **19**, 409 (1978).

19. G. D. Gladding, H. J. Kupferberg, and E. A. Swinyard, in *Handbook of Experimental Pharmacology*, Vol. 74, H. Frey and D. Janz, Eds., (Springer-Verlag, Berlin, 1985), p. 342.
20. J. T. Litchfield and F. J. Wilcoxon, J. Pharm. Exp. Ther. **96**, 99 (1949).
21. J. J. P. Stewart, in *Reviews in Computational Chemistry*, Vol. 1, K. B. Lipkowitz and D. B. Boyd, Eds. (VCH Publishers, New York, 1990).
22. J. J. P. Stewart, Mopac, version 7.0 F. J. Seiler Research Laboratory. United States Air Force Academy, CO 80840, 1994.
23. G. L. Estiú and L. Bruno-Blanch, Int. J. Quant. Chem.-Quant. Biol. Symp. **22**, 39 (1995).
24. M. C. Zerner, ZINDO Package, Quantum Theory Project, Williamson Hall, University Florida.
25. W. D. Edwards and M. C. Zerner, Theoret, Chim. Acta **72**, 347 (1987).
26. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, Gaussian 94, Revision B.3, Gaussian, Inc., Pittsburgh, PA, 1995.
27. G. L. Estiú, J. Mol. Structure (THEOCHEM), to appear.
28. R. S. Mülliken, J. Chem. Phys. **36**, 3428 (1962).
29. A. E. Reed, R. B. Weinstock, and F. Weinhold, J. Chem. Phys. **83**, 735 (1985). J. P. Foster and F. Weinhold, J. Am. Chem. Soc. **102**, 7211 (1980).
30. L. E. Chirlian and M. M. Francl, J. Comput. Chem. **8**, 894 (1987).
31. D. R. Lide, *CRC Handbook of Chemistry and Physics* (CRC Press: Boca Raton, FL, 1990–91).
32. K. B. Wiberg and K. E. Laidig, J. Am. Chem. Soc. **109**, 5935 (1987). K. B. Wiberg and C. M. Breneman, J. Am. Chem. Soc. **114**, 831 (1992).
33. R. E. Cachau, J. Mol. Biol. **255**, 31 (1996), Appendix II.
34. M. Karelson, V. S. Lobanov, and A, Katrizky, Chem. Rev. **96**, 1027 (1996).
35. A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry. Introduction to Advanced Electronic Structure* (McGraw-Hill, New York, 1989), p. 152.
36. Y-T. Chang, G. H. Loew, A. E. Rettie, T. A. Baillie, and P. R. Ortiz de Montellano, Int. J. Quant. Chem.-Quant. Biology Symp. **20**, 161 (1993).
37. S. Tasso, S-Ch. Moon, A. Goeta, G. Punte, G. L. Estiú, and L. Bruno-Blanch, submitted.
38. B. Macchia, A. Balasamo, M. C. Breschi, G. Chiellini, A. Lapucci, M. Macchia, C. Manera, A. Martinelli, C. Martini, R. Scatizzi, and U. Barreta, J. Med. Chem. **36**, 3077 (1993).
39. P. E. Keane, J. Simiand, E. Mendes, V. Santucci, and M. Moore, Neuropharmacology **22**, 875 (1983).

Conformational Analysis of Polypeptide Chains with the Aid of Density of States Calculations

SAUL G. JACCHIERI

Fundação Antônio Prudente, Rua Prof. Antônio Prudente 211, São Paulo, Brasil, SP 01509-900

Received 29 March 1997; revised 18 July 1997; accepted 11 August 1997

ABSTRACT: A model system whose density of states is an analytical function of the potential energy is obtained by combining potential energy wells given by Lennard-Jones 6-12 potentials representing pairwise interactions between atoms and circular barriers. Structural aspects of polypeptide chains such as sharp and broad energy extremes and close-packed and loose-packed conformations are simulated. By changing Lennard-Jones parameters, the density of states is described as a function of topological features of the potential energy surface and rules used to interpret density of states calculations are derived. Important results are that the number of clusters of density of states maxima in a given energy range approaches the number of conformational families and very low density of states gaps indicate the existence of kinetic barriers. These conclusions are applied to the conformational analysis of α-MSH. Structural implications are discussed.
© 1997 John Wiley & Sons, Inc. Int J Quant Chem **65**: 1115–1124, 1997

Key words: conformational analysis; density of states; potential energy surface

Introduction

The potential energy landscape (PEL) contains relevant information from which it could be possible, in principle, to extract all equilibrium and kinetic properties obtainable by theoretical means. However, a combinatorial explosion and/or the multiple minimum problem are faced in any attempt to carry out a detailed analysis of its topology. Although we may adopt a semiempirical potential energy function, it is not always possible to use it without further approximations. Recently, it was shown [1] that there are cases for which it is possible to find the global energy minimum. Even considering that the protein-folding problem can be solved when the absolute mini-

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061115-10

mum is reachable, the study of structure–function relationships presents a greater complexity, since, in many cases, the exercise of a function requires a structural transition, i.e., more than one energy minimum is involved.

The density of states or energy distribution may also be used to calculate equilibrium and kinetic properties, even though the interpretation of density of states calculations presents some ambiguities such as whether [2–4] a zero density of states gap indicates a folding tendency. Part of the difficulty is that whereas the semiempirical potential energy surface is an analytical function the density of states is, in most cases, a numerical function. In this article, we present a model that enables an analytical solution of the density of states as a function of the internal potential energy. We developed the idea, initially part of models used to study the dynamics of liquids [5], that the density of states obtained analytically or numerically as a function of the potential energy may be employed to investigate the existence of extremes in the potential energy landscape of a polypeptide chain within a chosen energy range.

The results obtained with the present model system are compared to calculations of a polypeptide chain. The 13-mer polypeptide chain of α-MSH was chosen as a suitable example. There are experimental and theoretical data related to the binding [6, 7] of α-MSH to biological membranes, indicating that α-MSH undergoes a structural transition as it moves from one chemical environment to the other.

A previous study [8] of the most probable conformational families adopted by α-MSH showed that a broad sampling of the conformational space necessary to describe a conformational transiton may be achieved with an unbiased conformational analysis in which no initial attempt is made to favor the occurrence of minimum-energy structures. Only after the determination of conformational families was an energy minimization procedure adopted to find stable conformations. This same example is used here to show how the results of a conformational analysis may be interpreted with density of states calculations.

Theory: Analytical Solution of a Model System

Even for the most simple dipeptide chain, it is not possible to derive the density of states η as an analytical function of the potential energy E. To overcome this limitation, we devised a model devoid of unnecessary complications that contains the essential characteristics that we want to consider.

Let us begin with this model, depicted in Figure 1, for which η is an analytical function of E. Our main assumption is that each atom interacts with barriers formed by the other atoms. That it should be so is justified by the consideration that in close-packed structures each atom is in the vicinity

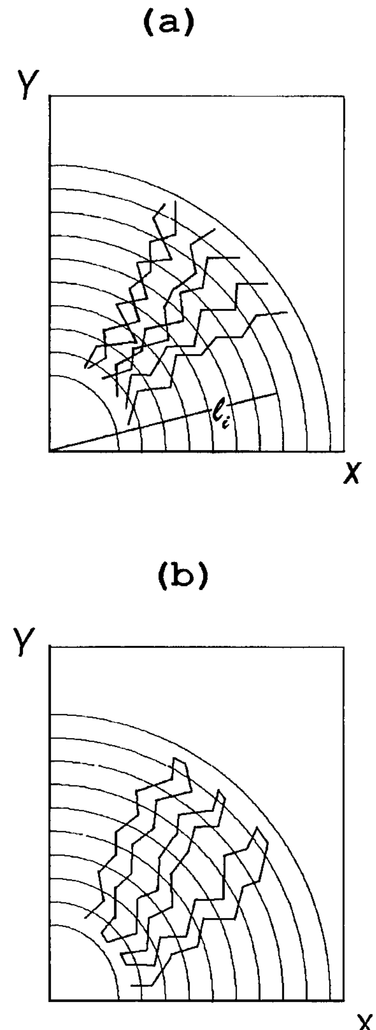


FIGURE 1. The model system. l_i are the radii of circular barriers as described in the Theory section. (a) Extended loose-packed conformational family generated with $l_1 < l_2 < \cdots < l_n$; (b) coiled close-packed conformational family generated with $l_{1+k} < l_{2+k} < \cdots < l_{n+k}$, $n = 1, 2, \ldots, k = 0, 1, 2, \ldots$.

of many surrounding atoms. On the other hand, when loose-packed flexible chains are being considered, the assumption of barriers also accounts for an averaging of atomic interactions in a large set of different conformations.

An important consequence of this approximation is that the potential energy may be considered as a function of distances between barriers and atoms $d_{ij} = l_j - \sqrt{x_i^2 + y_i^2}$ instead of interatomic distances $d_{ij} = \sqrt{x_{ij}^2 + y_{ij}^2}$. By making the further assumption that, as shown in Figure 1, the barriers are circles of radius l_i, we confer to the model system the circular symmetry that makes possible an analytical solution.

As shown in Figure 1(a), if the radii l_i follow an increasing order $(l_1 < l_2 < \cdots < l_N)$, the chain molecule is forced to adopt extended conformations. On the other hand, if as shown in Figure 1(b) the l_i have a periodicity similar to $l_{1+k} < l_{2+k} < \cdots < l_{n+K}$, $k = 0, 1, 2, \ldots$, a variety of coiled conformations becomes possible. Configurations of the model system generated with different sets of l_i's will be referred to as conformational families.

In the Appendix, the following equation is derived for the density of states $\eta(E)$ as a parametric function of the potential energy E:

$$\eta(E) = \sum_{i<j} S_{ij} \cdot \frac{l_j - H_{ij}}{6 \cdot H_{ij}^5} \cdot \left[\frac{b_{ij}}{2 \cdot V_{ij}^2} (-/+) \frac{\dfrac{b_{ij}^2}{2 \cdot V_{ij}^2} + \dfrac{a_{ij}}{V_{ij}}}{\sqrt{b_{ij}^2 + 4 \cdot a_{ij} \cdot V_{ij}}} \right], \quad (1)$$

where H_{ij} and S_{ij} are defined in Eqs. (A.6) and (A13), V_{ij} are 6-12 Lennard-Jones potentials [Eq. (A1)], a_{ij} and b_{ij} are the parameters of these potentials, and $E = \sum_{i<j} V_{ij}$. The minus and plus signs on the right-hand side apply, respectively, to positive and negative values of V_{ij}. Besides being a function of E, the density of states depends on the Lennard-Jones parameters of Eq. (1) and on the radii l_j of the circular barriers, as well.

If instead of this simple model used to derive Eq. (1) we started from a real polypeptide chain, we would need to make use of the equation

$$A = \int \int_D \cdots \int S(\chi_1, \chi_2, \ldots, \chi_N)\, d\chi_1\, d\chi_2 \cdots d\chi_N \quad (2)$$

to calculate the surface area of the potential energy landscape [S is given by Eq. (A.8), χ_i are internal degrees of freedom, and D is the domain encompassed by the χ_i]. The evaluation of the integrals in this equation is very difficult, if not impossible, and the density of states of a polypeptide chain is obtainable only as a numerical function of the potential energy surface. However, Eq. (2) shows that $\eta(E)$ may be seen as a projection of the potential energy landscape onto a two-dimensional space.

Methods

Five main-chain rotamers (A: $\phi = -57$, $\psi = -47$; B: $\phi = -139$, $\psi = 135$; G: $\phi = -60$, $\psi = -30$; D: $\phi = -90$, $\psi = 0$; and E: $\phi = 70$, $\psi = -60$) were employed in the search. For the side-chain conformations, we made use of the *gauche* minus, *trans*, and *gauche* plus rotamers of the Ponder and Richards classification (the numerals 1, 2, and 3 indicate, respectively, the *gauche* minus, *trans*, and *gauche* plus rotamers of χ_1) [9]. Gas-phase energies were calculated with the ECEPP/2 force field [10]. Hydration energies were calculated with the Ooi et al. [11] hydration potentials and the Connoly routine [12] for surface area computation. Matrix operations described in the literature [13] were employed in the conformational search with the following procedure: An initial set of conformations belonging to the amino end tetrameric fragment of the peptide chain is generated with a matrix operation. This set undergoes a selection in which those chain conformations populating density of states maxima and minima are selected, resulting in a much smaller set of conformations. Each of these latter chain conformations becomes then part of a matrix operation in which the structure of the amino end tetrameric fragment is frozen whereas the structure of the remaining part of the chain is varied. The resulting set of conformations is screened again following the same criterion of density of states maxima and minima. The same process is repeated until the end of the chain.

Results and Discussion

INTERPRETATION OF DENSITY OF STATES CALCULATIONS

By choosing appropriate values of the parameters in Eq. (1), it is possible to reproduce various

aspects of the structures adopted by polypeptide chains such as sharp and broad energy minima and close-packed and loose-packed conformations. A detailed examination of these calculations indicates how to find correspondences between features of potential energy surfaces and density of states plots.

A comparison between the density of states and the potential energy surface may be carried out for one (i.e., for one set of l_i's) or many conformational families. The most important features that we want to consider are the number and location of density of states maxima and minima and their relative magnitudes, compared to the location of extremes in the potential energy surface.

The potential energy surface given by Eq. (A.3) is rugged, since it increases abruptly whenever an atom approaches a barrier. On the other hand, the single atom interaction energy V_{ij} is smooth and has a minimum when the distance (d_{ij}) between atom i and barrier l_j is $[l_j - 2 \cdot a_{ij}/b_{ij}]^{1/6}$. To have a smooth representation of the potential energy surface, we have defined a parameter χ so that

$$d_{ij} = \frac{\left(l_j - \dfrac{2 \cdot a_{ij}}{b_{ij}}\right)^{1/6}}{F} \cdot \chi,$$

and V_{ij} is a minimum when $\chi = F$. The potential energy E is plotted as a function of χ considering that all d_{ij}'s change in concert with $\chi = 0$ to $\chi = \infty$.

In Figure 2, the coexistence of two different conformational families is illustrated. As shown in Figure 2(b) and (c), low-energy density of states maxima are populated mainly by the lowest-energy conformations of each conformational family. According to these results, clusters of density of states maxima separated by very low density of states gaps may be interpreted as distinct conformational families, and a gap in η indicates the presence of a kinetic barrier [as shown in Eqs. (A.15) and (A.16)]. This same analysis shows, however, that if the minimum-energy conformations of different conformational families have similar energies the corresponding clusters of density of states maxima may overlap. In the example depicted in Figure 2(d), there is a partial overlap of density of states maxima.

In Figure 3, the model system is allowed to occupy energy wells corresponding to the configurations represented in Figure 1(a) and (b). The lowest potential energy well corresponds to the coiled and close-packed configuration [Fig. 1(b)], whereas the highest potential energy well (which is also broader) shown in Figure 3(a) corresponds to the extended and loose-packed configuration [Fig. 1(a)]. A comparison between Figure 3(b) and (c) shows that the loose-packed conformational family has a much larger density of states, as one would expect intuitively. The sudden increase of $\eta(E)$ seen in Figure 3(d) when the energy increases above the threshold where the transition from coiled to extended conformational families takes place is seen in the present (see Fig. 4) and in previous calculations [2–4, 14, 15].

The above-described analysis shows that $\eta \times E$ plots have an inherent ambiguity. Density of states gaps should correspond to a folding tendency if there are clusters of density of states maxima separated by these gaps populated by unique conformational families, and in this case, it also indicates the existence of a kinetic barrier [see Eqs. (A.14)–(A.16) of the Appendix for details]. However, the same conformational family may populate more than one cluster of density of states maxima in which case a folding tendency should not exist.

Usually, it is not possible to tell the two situations apart just by examining $\eta \times E$ plots and some other source of evidence is necessary. As shown in the next section, by combining density of states calculations with a molecular graphics study of the conformations that belong to different regions of the $\eta \times E$ plot, we may execute a detailed conformational analysis.

APPLICATION TO POLYPEPTIDE CHAINS

The same basic principles obtained from the above examples are considered to be valid for real systems. Our main purposes were to identify conformational families, to evaluate their relative stabilities, and to determine the presence of kinetic barriers between structural transitions. This is achieved by generating $\eta \times E$ plots as those shown in Figure 4.

Details of the procedure used in these calculations are discussed in the Methods section and in [13]. In principle, any algorithm that fulfills the requirement of an unbiased conformational search may be used to calculate the density of states. In this article, we used the matrix algorithm [13] for this purpose.

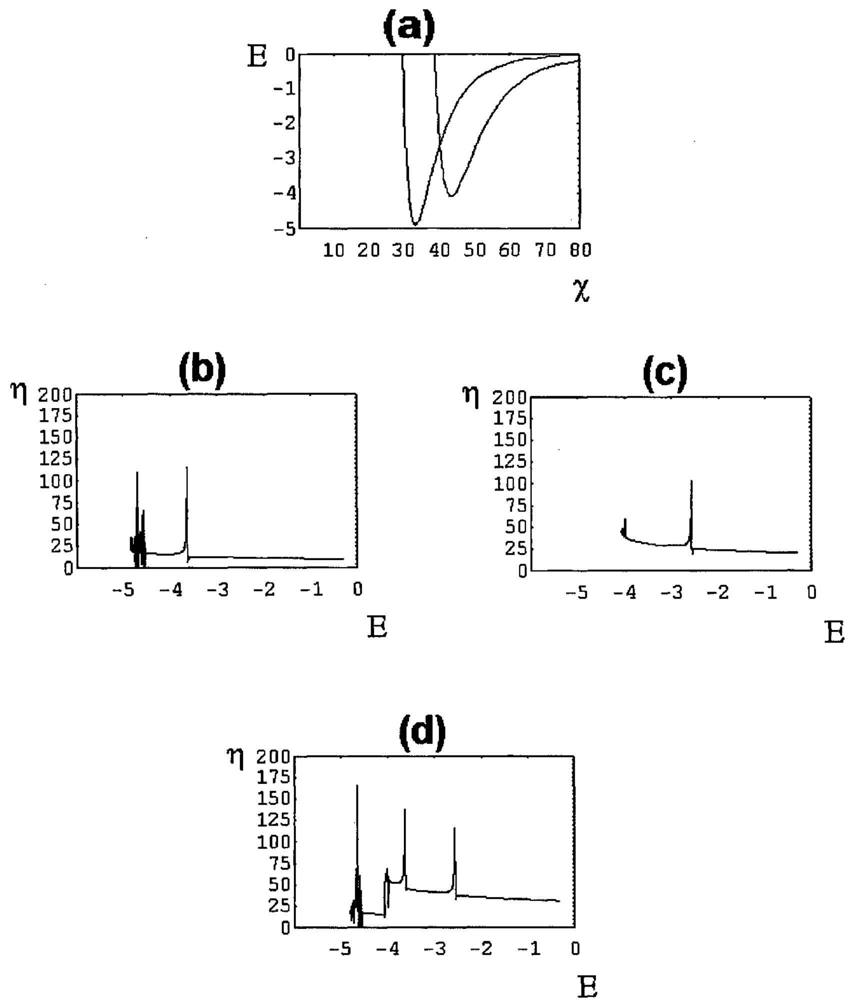


FIGURE 2. (a) Potential energy wells of two extended configurations of the model system. $\eta \times E$ plots are shown: (b) the lower-energy well corresponds to a configuration having five atoms, $a_{ij} = 0.1$, $b_{ij} = 1$, and $l_1 < l_2 < \cdots < l_5$; (c) corresponds to an extended configuration having the same number of atoms and same Lennard-Jones parameters and radii $m_i = l_i + 1$; (d) $\eta \times E$ plot showing a sum of the $\eta(E)$ values belonging to the described conformational families.

The existence of three α-MSH populations, or conformational families, in aqueous and lipidic environments is supported by theoretical and experimental [6, 7] evidence. This result has been interpreted [6–8] as being directly related to the three rotamer conformations of χ_1 of Trp 9.

In Figure 4, we also show the energies for which the most stable conformations belonging to each of the three rotamers of χ_1 of Trp 9 were found. The *trans* rotamer of Trp 9 is the first density of states minimum, whereas the *gauche* minus and *gauche* plus rotamers are just after the second density of states minimum.

The existence of a density of states minimum separating two maxima indicates a folding tendency [2–4] if each maximum (or, presumably, the lowest-energy maximum) is populated by only one conformational family. As shown in Table I, the lowest-energy density of states maximum is populated only by conformations adopting the *trans* rotamer of χ_1 for the side-chain rotamer of Trp 9. However, regions (b) and (c) belonging to the second density of states maximum are, respectively, populated by *gauche* minus and *trans* and by *gauche* plus and *trans* rotamer conformations of Trp 9.

These results indicate that α-MSH has a folding tendency favoring the *trans* rotamer of Trp 9, even though it is not possible to point to a kinetic barrier because there is no zero (or very small

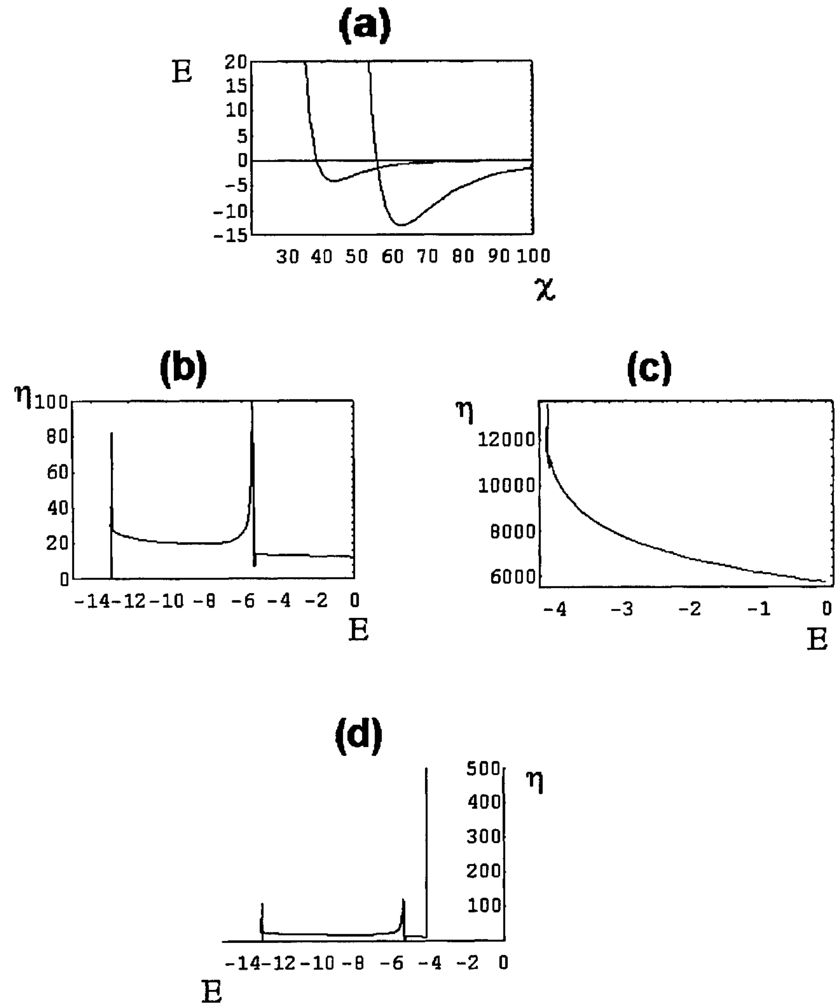


FIGURE 3. (a) Potential energy wells of two configurations of the model system. $\eta \times E$ plots are shown: (b) the lower-energy well corresponds to a coiled configuration having nine atoms and $a_{ij} = 0.1$ and $b_{ij} = 1$ and a turn at every three atoms, whereas the higher-energy well (c) corresponds to an extended configuration having the same number of atoms and same Lennard-Jones parameters; (d) $\eta \times E$ plot showing a sum of the $\eta(E)$ values belonging to the described conformational families.

density) of states minimum. A tendency to one conformation or conformational family might change if α-MSH were allowed to coexist in the aqueous solution and a lipidic phase. To examine these questions, we calculated the hydration energies of all conformations within the energy range delimited by the energy intervals (a) and (c) of Figure 4(i). As shown in Figure 4(ii), the energy difference between the energy intervals in which chain conformations (a) and (b) [and (c)] are found increased by 5 kcal/mol when hydration energies are considered, an indication that the *trans* rotamer of Trp 9 is stabilized in an aqueous solution.

In Figure 5, we show the most stable conformations belonging to the three populations discussed above. It is seen that transitions between the rotamer conformations of Trp 9 side chains are associated with extensive structural transitions encompassing the entire chain.

Conclusion

One of the main advantages of the application of density of states calculations to the conformational analysis of polypeptide chains is that no matter how long is the chain the density of states is always a function of the internal energy and

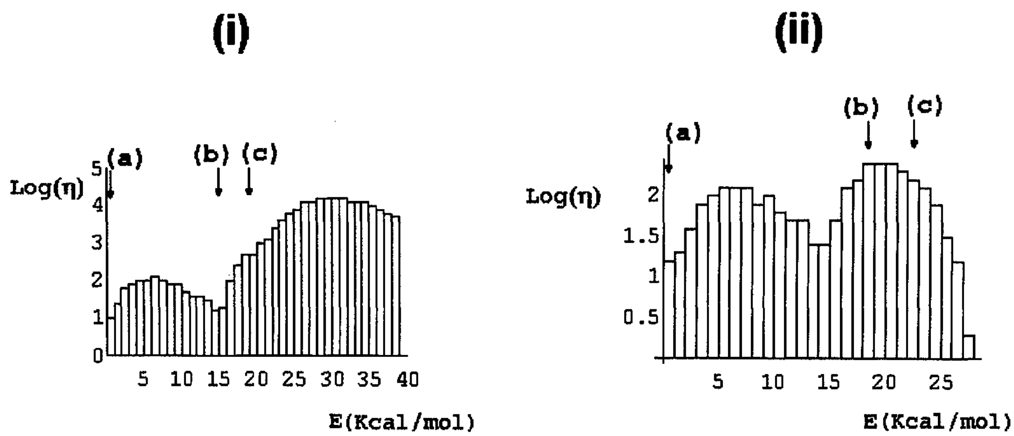


FIGURE 4. $\log(\eta) \times E$ plots calculated for α-MSH with (i) ECEPP/2 gas-phase potentials and (ii) ECEPP/2 [11] gas-phase potentials plus hydration [12] potentials. (a), (b), and (c) are the energy intervals where the most stable conformations adopting, respectively, the *trans*, *gauche* minus, and *gauche* plus conformers of Trp 9. The energy E is rescaled so that the minimum-energy conformation is the zero-energy level.

$\eta \times E$, a two-dimensional graph, and this is a consequence of the density of states being a projection of the potential energy landscape.

However, the integrals of Eq. (2) cause a loss of information and the larger the number of internal degrees of freedom the more information is lost. The only information maintained is the one mostly needed in the study of structure and function relationships, e.g., the number of conformational families in the considered energy range and their relative stabilities.

An important consequence of Eqs. (A.15) and (A.16) is that to every local energy minimum corresponds a density of states maximum. Density of

TABLE I

Ten more stable conformations belonging to the energy intervals indicated as (a), (b), and (c) in Figure 4 which correspond to the *trans*, *gauche* minus, and *gauche* plus conformations of χ_1 of Trp 9; the same conformations are depicted in Figure 5.

(a)		(b)		(c)	
GAAAAAGDDDBGB[a] 1121122121312	−0.4645100E + 05[b]	DDAGDGAAAABBB 1111222211111	−0.3038100E + 05	DDGBDGBBGABBB 2121322131312	−0.2591100E + 05
GAAAAAGDDDBGB 1121122121313	−0.4642300E + 05	DDAGDGAAAABGB 1111222211112	−0.3029900E + 05	GDDBBDAAAABDB 1111112221113	−0.2591000E + 05
GAAAAAGDDDBGB 1121122121311	−0.4602900E + 05	BGGGDAAAAABBB 1212221221313	−0.3029200E + 05	BGGGDAAAAABBA 1212221221311	−0.2591000E + 05
GAAAAAGDDDBAG 1121122121313	−0.4587100E + 05	BGGGDAAAAABGG 1212221221312	−0.3028900E + 05	ADABGAAAAABBA 3331112221111	−0.2591000E + 05
GAAAAAGDDDBAA 1121122121313	−0.4570200E + 05	DDAGDGAAAABGB 1111222211113	−0.3025200E + 05	ADDBBDAAAABDB 1111112221113	−0.2590900E + 05
GAAAAAGDDDBBB 1121122121312	−0.4545100E + 05	ADABGAAAAABAG 3331112221113	−0.3021800E + 05	GDDBBDAAGABGB 1111112221313	−0.2590800E + 05
GAAAAAGDDDBGG 1121122121312	−0.4543300E + 05	BGGGDAAAAABAB 1212221221312	−0.3021500E + 05	BGABBDGAAABBG 1231322221112	−0.2590800E + 05
GAAAAAGDDDBAB 1121122121312	−0.4533500E + 05	ADABGAAAAABBB 3331112221112	−0.3021000E + 05	ADDBBDAAGABGB 1111112221313	−0.2590600E + 05
GAAAAAGDDDBAB 1121122121313	−0.4530400E + 05	ADABGAAAAABBG 3331112221113	−0.3019700E + 05	ADBBBBGAAABAA 2131112121113	−0.2590600E + 05
GAAAAAGDDDBGA 1121122121312	−0.4510200E + 05	BGGGDAAAAABBB 1212221221112	−0.3018500E + 05	DDAGDGAAAABEG 1111222211212	−0.2590500E + 05

[a] Letters and numerals indicate, respectively, monomer residue main-chain and side-chain conformers as defined in the Methods section.
[b] ECEPP/2 gas-phase energy (kcal/mol).

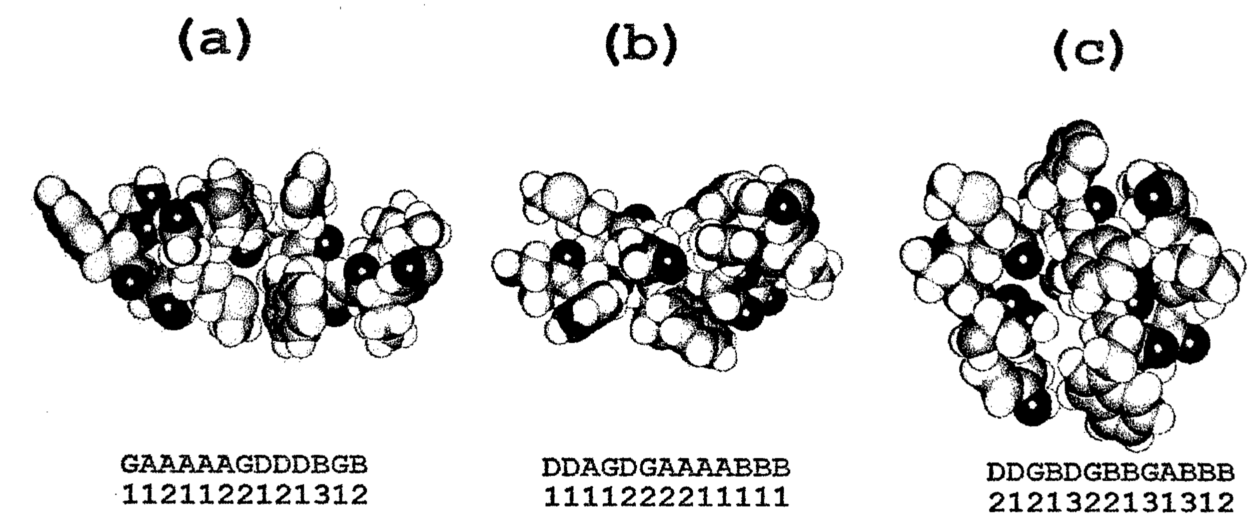


FIGURE 5. Most stable conformations of α-MSH found in the present conformational search adopting the (a) *trans*, (b) *gauche* minus, and (c) *gauche* plus conformations of χ_1 of Trp 9. (1) Letters and numerals indicated main-side and side-chain rotamers (see Methods section for a definition). In Figure 4, the corresponding energy intervals are indicated by the arrows. Color coding: white: hydrogen; black: carbon; blue: nitrogen; red: oxygen; yellow: sulfur.

states maxima belonging to close energy minima clusterize and become indistinguishable. We can still postulate, however, that the number of clusters of the density of states maxima is less or equal to the number of conformational families.

A similarity between the presently shown results and previous calculations is seen, for instance, in conformational searches [15] with a lattice model and reduced energy representations. As in the present treatment, conformations are generated regardless of their energies and subsequently classified in the order of increasing energy. In all cases, the native structure belongs to a small maximum in the low-energy side of energy distributions. Another example is the energy distribution of normal modes of vibration [14] which shows that near- and far-from-equilibrium states belong to distinct regions of density of states vs. energy plots.

Appendix: The Model System Density of States vs. Energy Plot

The pairwise molecular mechanics atomic interactions between atoms i and j are functions of the interatomic distances $d_{ij} = \sqrt{x_{ij}^2 + y_{ij}^2}$. However, in the model system depicted in Figure 1, we consider the interaction between atom i and circular barriers formed by the remaining atoms. This approximation confers to our model the circular symmetry that makes possible an analytical solution of the problem of describing the density of states as a function of the potential energy surface. The interatomic interactions are replaced by interactions between atoms and circular barriers of radius l_j. The proper distances are

$$d_{ij} = l_j - \sqrt{(x_i^2 + y_i^2)}\,. \tag{A.1}$$

By applying this approximation to all pairwise interactions, the nonbonded contribution to the potential energy becomes

$$E(x_1, y_1, x_2, y_2, \ldots) = \sum_{i<j} V_{ij}, \tag{A.2}$$

where

$$V_{ij} = \frac{a_{ij}}{\left(l_j - \sqrt{(x_i^2 + y_i^2)}\right)^{12}} - \frac{b_{ij}}{\left(l_j - \sqrt{(x_i^2 + y_i^2)}\right)^{6}} \tag{A.3}$$

and a_{ij} and b_{ij} are Lennard-Jones parameters.

In Figure 1, it is shown how different extended and coiled conformations may be generated with a proper choice of the radii l_j. As shown in the following derivation, the density of states becomes a function of the Lennard-Jones force-field parame-

ters a_{ij} and b_{ij} and of the structural parameters l_j, as well.

Let us consider the region of the potential energy surface delimited by the conditions

$$0 \le x_i \le l_i \quad (A.4)$$

$$\sqrt{l_i^2 - x_i^2} \le y_i \le \sqrt{(l_j - H_{ij})^2 - x_i^2}, \quad (A.5)$$

where

$$H_{ij} = \left(\frac{-b_{ij}(+/-)\sqrt{b_{ij}^2 + 4 \cdot a_{ij} \cdot V_{ij}}}{2 \cdot V_{ij}} \right)^{1/6}. \quad (A.6)$$

Within this region, the potential energy given by Eq. (A.3) has a lower value of E.

The number of states that may be occupied by atom i within the boundaries defined by Eqs. (A.4) and (A.5) is proportional to the surface area:

$$A_i = \sum_{j>i} \int_0^{l_i} \int_{\sqrt{l_i^2 - x_i^2}}^{\sqrt{(l_j - H_{ij})^2 - x_i^2}} S_{ij}(x_i, y_i)\, dx_i\, dy_i, \quad (A.7)$$

where

$$S_{ij}(x_i, y_i) = \sqrt{1 + \left(\frac{\partial V_{ij}}{\partial x_i} \right)^2 + \left(\frac{\partial V_{ij}}{\partial y_i} \right)^2}. \quad (A.8)$$

The total number of conformations available to the model system within this region is then proportional to the total surface area $A = \sum_i A_i$ and the density of states is given by

$$\eta(E) = \frac{\partial A}{\partial E} = \sum_{i<j} \frac{\dfrac{\partial A}{\partial V_{ij}}}{\dfrac{\partial E}{\partial V_{ij}}}. \quad (A.9)$$

Replacing this relation in (A.7) and applying the Leibnitz formula [16], we obtain

$$\eta(E) = \sum_{i<j} \int_0^{l_i} S_i(x_i, I_{ij}(x_i, V_{ij})) \cdot \frac{\partial I_{ij}(x_i, V_{ij})}{\partial V_{ij}} \cdot dx, \quad (A.10)$$

where

$$I_{ij}(x_i, V_{ij}) = \sqrt{(l_j - H_{ij})^2 - x_i^2}. \quad (A.11)$$

In this simple instance, it is possible to integrate Eq. (A.10) analytically:

$$\eta(E) = \sum_{i<j} S_{ij}(V_{ij}) \cdot \frac{l_j - H_{ij}}{6 \cdot H_{ij}^5} \cdot \left[\frac{b_{ij}}{2 \cdot V_{ij}^2}(-/+)\frac{\dfrac{b_{ij}^2}{2 \cdot V_{ij}^2} + \dfrac{a_{ij}}{V_{ij}}}{\sqrt{b_{ij}^2 + 4 \cdot a_{ij} \cdot V_{ij}}} \right], \quad (A.12)$$

where

$$S_{ij} = \sqrt{1 + \frac{1}{(l_j - H_{ij})^4} \left(\frac{6 \cdot b_{ij}}{H_{ij}^7} - \frac{12 \cdot a_{ij}}{H_{ij}^{13}} \right)} \quad (A.13)$$

is obtained by replacing Eq. (A.11) in the term $S_{ij}(x_i, I_{ij}(x_i, V_{ij}))$ of Eq. (A.10).

$\eta \times E$ plots may then be obtained for various values of the parameters in Eq. (A.12). A numerical analysis of $\eta(E)$ showed that the behavior of the first term on the right-hand side of Eq. (A.12) is dominated by the expressions

$$T(V_{ij}) = \frac{b_{ij}}{2 \cdot V_{ij}^2}(-/+)\frac{\dfrac{b_{ij}^2}{2 \cdot V_{ij}^2} + \dfrac{a_{ij}}{V_{ij}}}{\sqrt{b_{ij}^2 + 4 \cdot a_{ij} \cdot V_{ij}}}. \quad (A.14)$$

$T(V_{ij})$ obeys the following limits, when the second term on the right-hand side is subtracted from the first:

$$\lim_{V_{ij} \to -\frac{b_{ij}^2}{4a_{ij}}} T(V_{ij}) = -\infty$$
$$\lim_{V_{ij} \to 0} T(V_{ij}) = 0 \quad (A.15)$$
$$\lim_{V_{ij} \to \infty} T(V_{ij}) = 0$$

and when it is added

$$\lim_{V_{ij} \to -\frac{b_{ij}^2}{4a_{ij}}} T(V_{ij}) = \infty$$
$$\lim_{V_{ij} \to 0} T(V_{ij}) = \infty \qquad \text{(A.16)}$$
$$\lim_{V_{ij} \to \infty} T(V_{ij}) = 0.$$

$b_{ij}^2/4 \cdot a_{ij}$ is (for large l_j) the value of the energy minimum of V_{ij}.

ACKNOWLEDGMENTS

We gratefully acknowledge time provided by the NCI's Frederick Biomedical Supercomputer Center and support provided by staff personnel.

References

1. H. A. Scheraga, Biophys. Chem. **59**, 329 (1996).
2. M.-H. Hao and H. A. Scheraga, J. Chem. Phys. **102**, 1334 (1995).
3. A. Sali, E. Shakhnovich, and M. Karplus, J. Mol. Biol. **235**, 1614 (1994).
4. M. Fukugita, D. Lancaster, and M. G. Mitchard, Biopolymers **41**, 239 (1997).
5. B. Madan and T. Keyes, J. Chem. Phys. **98**, 3342 (1993).
6. A. S. Ito, A. M. L. Castrucci, V. J. Hruby, M. E. Hadley, D. T. Krajcarski, and A. G. Szabo, Biochemistry **32**, 12264 (1993).
7. F. Al Obeidi, M. E. Hardley, B. M. Pettit, and V. J. Hruby, J. Am. Chem. Soc. **111**, 3413 (1989).
8. S. G. Jacchieri and A. S. Ito, Int. J. Quantum Chem. **53**, 335 (1995).
9. J. W. Ponder and F. M. Richards, J. Mol. Biol **193**, 775 (1987).
10. S. S. Zimmerman, M. S. Pottle, G. Nemethy, and H. A. Scheraga Macromolecules **10**, 1 (1977).
11. T. Ooi, M. Oobatake, G. Nemethy, and H. A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **84**, 3080 (1987).
12. M. L. Connoly, J. Appl. Crystallogr. **16**, 548 (1983).
13. S. G. Jacchieri and R. L. Jernigan, Biopolymers **32**, 1327 (1992).
14. J. E. Straub and Ji-Kyung Choi, J. Phys. Chem. **98**, 10978 (1994).
15. D. J. Covell and R. L. Jernigan, Biochemistry **29**, 3287 (1990).
16. M. Abramovitz and I. Stegun, Eds., *Handbook of Mathematical Functions* (Dover, New York, 1992), p. 11.

A Semiempirical Study on Leupeptin: An Inhibitor of Cysteine Proteases

GABRIELA BARREIRO,[1] RICARDO BICCA DE ALENCASTRO,[1] JOAQUIM DELPHINO DA MOTTA NETO[2]

[1]*Physical Organic Chemistry Group, Departamento de Química Orgânica, Instituto de Química da UFRJ, Cidade Universitária, CT, Bloco A, lab. A609, Rio de Janeiro, RJ 21949-900, Brazil*
[2]*348 Williamson Hall, Quantum Theory Project, P.O. Box 118435, University of Florida, Gainesville, Florida 32611-8435*

Received 28 February 1997; revised 16 July 1997; accepted 15 August 1997

ABSTRACT: In this work, we modeled leupeptin (Ac.Leu.Leu.Arg.CHO), a natural inhibitor of proteases, and the active site of papain, a cysteine protease, using as a template the crystal structure of a leupeptin–papain complex recently obtained by Schroder and co-workers [FEBS Lett. **135**(1), 38 (1993)] and including 11 amino acids relevant to the proteolytic activity of the enzyme. Our results show that the AM1 fully optimized leupeptin is more stable than is the leupeptin crystal structure by about 6.0 kcal/mol. Our results show also that in the modeled active center of papain the S—H ··· N structure is favored. When the aldehyde is included in the calculation, however, proton transfer occurs with a strengthening of the $S^- \cdots HIm^+ \cdots O{=}C$ (Asn175) catalytic triad. The AM1 method reproduces fairly well the interactions between the enzyme and the host molecule. © 1997 John Wiley & Sons, Inc. Int J Quant Chem **65**: 1125–1134, 1997

Key words: leupeptin; papain; semiempirical; cysteine protease; active center

Introduction

Leupeptin is an acyltripeptide aldehyde (Ac.Leu.Leu.Arg.CHO) of mycrobial origin that has been used to investigate the possible role of protease in several important biological processes [1–3]. Leupeptin anticancer activity is poorly understood and even less is known about its steroid binding inhibition activity [4, 5]. Leupeptin inhibits serine proteases such as plasmin and trypsin, as well as cysteine proteases, papain among others [6, 7]. A few years ago, Schröder and co-workers obtained the X-ray crystal structure of the leupeptin–papain complex at 2.1 Å resolution [8] and, very recently, Kurinov and Harrison reported X-ray structures for two different crystal

Correspondence to: R. Bicca de Alencastro.
Contract grant sponsors: CNPq (Brazil); FAPERJ (Brazil).

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061125-10

forms of the leupeptin–trypsin complex at 1.7 Å resolution [5].

It is a great concern that malaria, a disease responsible for millions of deaths per year in tropical areas of the world, is consistently spreading, the main cause of which is the increasing resistance of *Plasmodium falciparum* to conventional drug therapy [9–13]. Proteases, on the one hand, are critical in many steps of the biological cycle of parasites and this makes them attractive for use in the development of new drugs [11, 14–18]. Cysteine proteases, on the other hand, have been used in last years as promising targets to new therapies (see [19] and references therein). Papain from *Candida papaya*, a representative of a superfamily that is predominant in eukariots, the active center of which is highly conserved [17], is attractive for study because its three-dimensional structure is well known [20–23]. The catalytic mechanism of this enzyme, first proposed by Drenth and co-workers [20, 21], has been the object of extensive experimental [24, 25] and theoretical work [26–35].

In the present article, we discuss are AM1 results on leupeptin and on a model of the leupeptin–papain complex, including 11 amino acids relevant to the active center of the enzyme. Our aim was to better understand some of the details of complex formation, having in mind the modeling of new antimalarials.

Method and Computational Details

All calculations were carried out at the SCF level using the AM1 [36–38] semiempirical Hamiltonian within the MOPAC 7.0 program [39] on IBM RISC 6000 workstations or within the UNICHEM package of a Cray J90 (NACAD-COPPE-UFRJ). The coordinates of the leupeptin–papain complex [8] were obtained from the Protein Data Bank.

Results and Discussion

The cysteine proteases are a group of proteolytic enzymes whose activity depends on a free thiol group of a cysteine residue [40]. The best-known cysteine proteases are obtained from plant sources. Among those, papain has become a model enzyme in enzymology [22, 23, 41] and in structure–activity studies [26–35]. We briefly described the main features of the active site of papain, as viewed by X-ray crystallography and enzymology studies (see [19] and references therein; for an extensive discussion of the catalytic mechanism of cysteine peptidases, see [22]). Unfortunately, X-ray structures of proteins are often obtained either in an inactive form, as, e.g., the structure of papain published by Kamphuis and co-workers (in which the thiol group of Cys25 is oxidized to SO_3^-) [42], or in an inhibited form, as in protein–inhibitor complexes [5, 8]. On the other hand, theoretical treatment of the problem is hindered by the number of atoms that must be included. For this reason, accurate methods of calculation have often been applied to very inaccurate models or, at the other extreme, force fields have been used for docking purposes. In this work, we present our calculations at the AM1 level of both the active center of papain and a papain–leupeptin complex, including a significant number of amino acids relevant to the proteolytic activity of the enzyme. We have not sought to include solvent molecules as yet, and therefore our results must be viewed as preliminary.

LEUPEPTIN

Leupeptin (Fig. 1) exists in three covalent forms in water at room temperature: leupeptin hydrate (42%), a cyclic carbinolamine (56%), and the free aldehyde (2%). The free aldehyde binds to papain with a binding constant equal to 2.2×10^{-11} *M* and is the sole active form of leupeptin [43]. In the free aldehyde form, leupeptin is a very flexible molecule. We calculated AM1 partial potential energy surfaces (PES) for 21 torsional angles (Fig. 1), keeping constant at 180° the three peptide bonds only. After that, starting from a geometry that retained the best values of the 21 torsional angles, we fully optimized the structure (Fig. 2). The results for this last geometry are shown in Table I. Finally, each one of the structures generated was superimposed on the crystal structure of leupeptin from the coordinates available in the Protein Data Bank [8]. The root mean-square values of the superimposed structures lie between 1.95 and 2.09 Å, except for the fully optimized geometry (Table I), for which it is 1.67 Å. Therefore, our best calculated structure ($\Delta H_f = -65.8$ kcal/mol) has a close fit to the crystal structure (Fig. 3, Table I). In an attempt to improve the fitting, we rotated torsional angles 3, 7, and 19 by 180° and fully optimized the structure so obtained. The resulting structure was

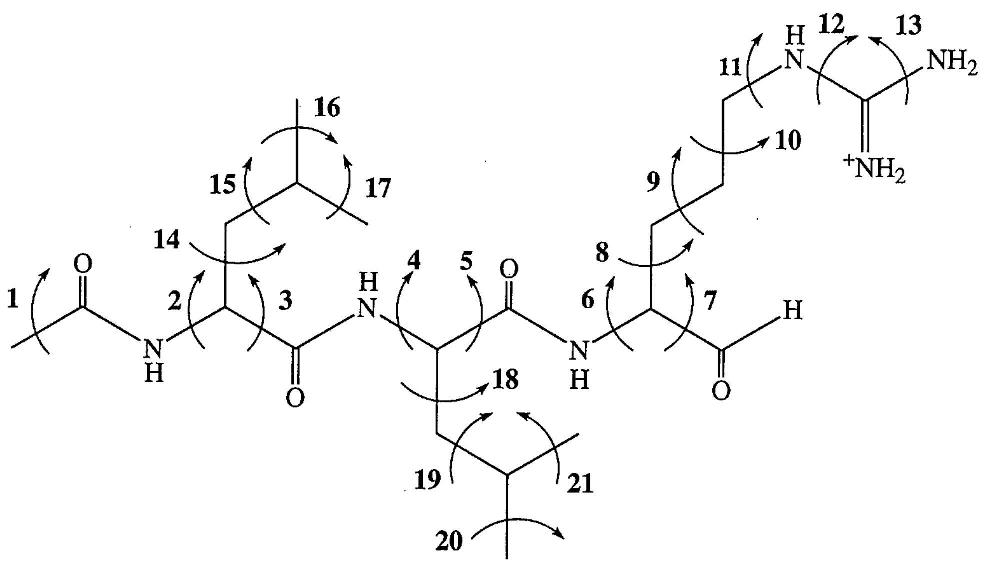


FIGURE 1. Leupeptin structure and optimized dihedral angles.

less stable than our calculated structure by $\Delta H =$ 7.4 kcal/mol (Table I). We also fully optimized the geometry of leupeptin obtained directly from the crystal structure. The result was a third structure 6 kcal/mol less stable than our calculated structure. These results seem to imply that in the crystal form of the leupeptin–papain complex the host molecule is in a tensioned conformation. By thermodynamic reasons, formation of the thiohemiacetal alone cannot explain this conformation [44]; therefore, stabilization must come from hydrogen bonds and other polar interactions, either from the water (and methanol) molecules or the peptide backbone [41].

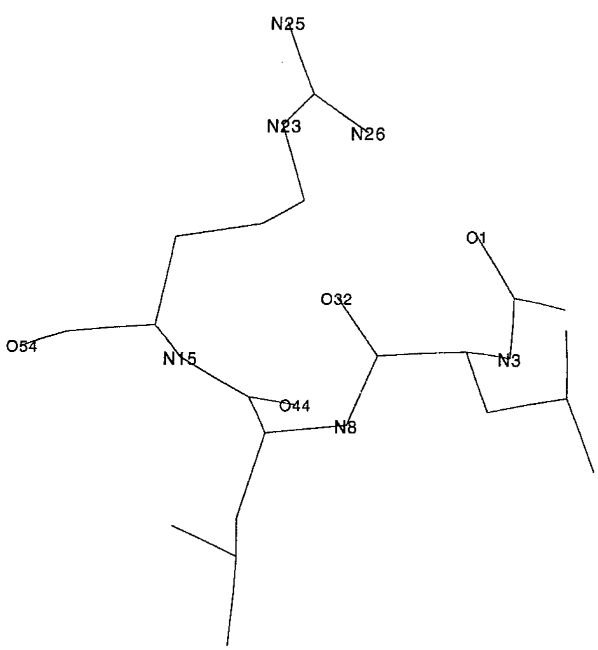


FIGURE 2. Optimized (AM1) leupeptin geometry.

LEUPEPTIN–PAPAIN COMPLEX

We now describe our model of the active site of papain and of the papain–leupeptin complex. Due to the computational effort required, we limited the model to leupeptin and 11 amino acids relevant to the active center. From the L-domain of papain (Fig. 4), we chose to include the following residues: Gly19, which contributes to the stabilization of the oxyanion hole of the enzyme [45] and forms hydrogen bonds to the α-helix [8]; the triad Gly23, Ser24, Cys25, which is the section of the α-helix involved in the oxyanion stabilization and the proteolytic activity [22]. We hope that we have included most of the electrostatic contribution of the α-helix to the stability of the catalytic triad [26, 30]. We also included Gly66, which is relevant to the binding of the host peptide during proteolysis [22], and Tyr67 and Pro68, two amino acids from the S_2 subsite (hydrophobic pocket; for nomenclature see [46]). From the R-domain of papain, we included Ala160, which is relevant for future studies concerning the development of antimalarials [19]; Asp158, which has been thought to contribute to the stability of the catalytic triad [23] (see,

TABLE I
Summary of dihedral angles of leupeptin.

	I AM1-optimized $\Delta H_f = -65.86$ (kcal / mol)	II Crystal structure (from PDB [8]) $\Delta H_f = 48.21$ (kcal / mol)	III Modified (from I)[a] $\Delta H_f = -58.45$ (kcal / mol)	IV AM1-optimized (from II) $\Delta H_f = -59.73$ (kcal / mol)
1	56.0	60.7	59.1	63.7
2	161.8	160.3	152.7	151.4
3[b]	114.4	−38.4	−111.5	−105.1
4	160.8	134.8	171.7	157.1
5	−37.9	−33.8	−28.6	5.0
6	96.0	118.0	95.2	106.4
7[b]	−17.4	168.2	−16.3	−159.5
8	−69.5	−90.1	−71.6	−120.1
9	96.4	118.8	95.2	62.3
10	68.4	68.9	66.6	63.7
11	−178.1	−130.9	−175.8	−152.5
12	−4.5	−0.2	−8.0	−8.2
13	6.6	2.2	15.9	−3.3
14	−71.5	−100.8	−66.2	−149.8
15	163.1	178.7	166.1	−159.1
16	−179.6	−180.0	−178.1	−180.1
17	179.1	−180.0	−180.0	−178.1
18	−67.3	−65.2	−69.1	−110.5
19[b]	−72.3	−46.8	−52.9	−67.1
20	179.7	178.0	−168.8	−177.8
21	−178.3	179.0	159.4	171.7

[a] Starting from structure **I**, these angles have been changed by 180° and the new geometry was fully optimized to reproduce the crystal structure (**II**).

however, [22]); and His159 and Asn175, directly involved in the proteolytic activity [22].

Table II shows the dihedral angles relevant for our models. Small differences can be observed between the dihedral angles in the peptide sections of the active site obtained from the crystal structure (column 1) and the AM1-optimized active site (after exclusion of leupeptin and water molecules; column 2). Most of the differences in the backbone dihedral angles must be ascribed to the energy optimization of the terminal groups of the fragments. Some differences, however, are relevant to the mechanism of enzymatic action. Figure 5 shows that exclusion of the host molecule (and water) leads in the AM1-calculated active site to deprotonation of the imidazole ring. This is followed by a torsion of the —CH_2SH away from the imidazole ring of His159. A similar torsion of the terminal —NCH_3 group of the Cys25 fragment approaches this group to —CH_2SH. However, it is possible that, at last partially, the rotation of the —CH_2SH group is an artifact of the calculation. Another potentially relevant difference is a reorientation of the OH lateral chain of Ser24 toward —CH_2SH, a residue hitherto not implied in the

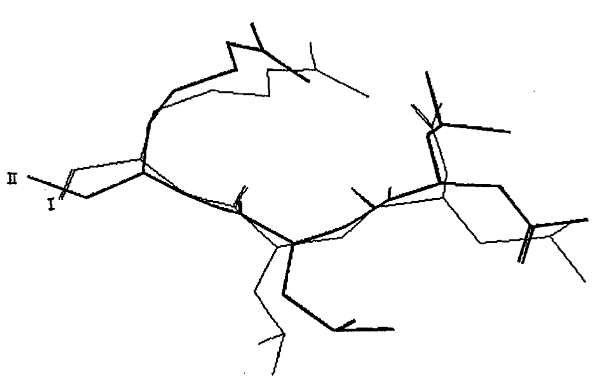


FIGURE 3. Superimposition of the optimized (AM1) and the X-ray structure geometries of leupeptin.

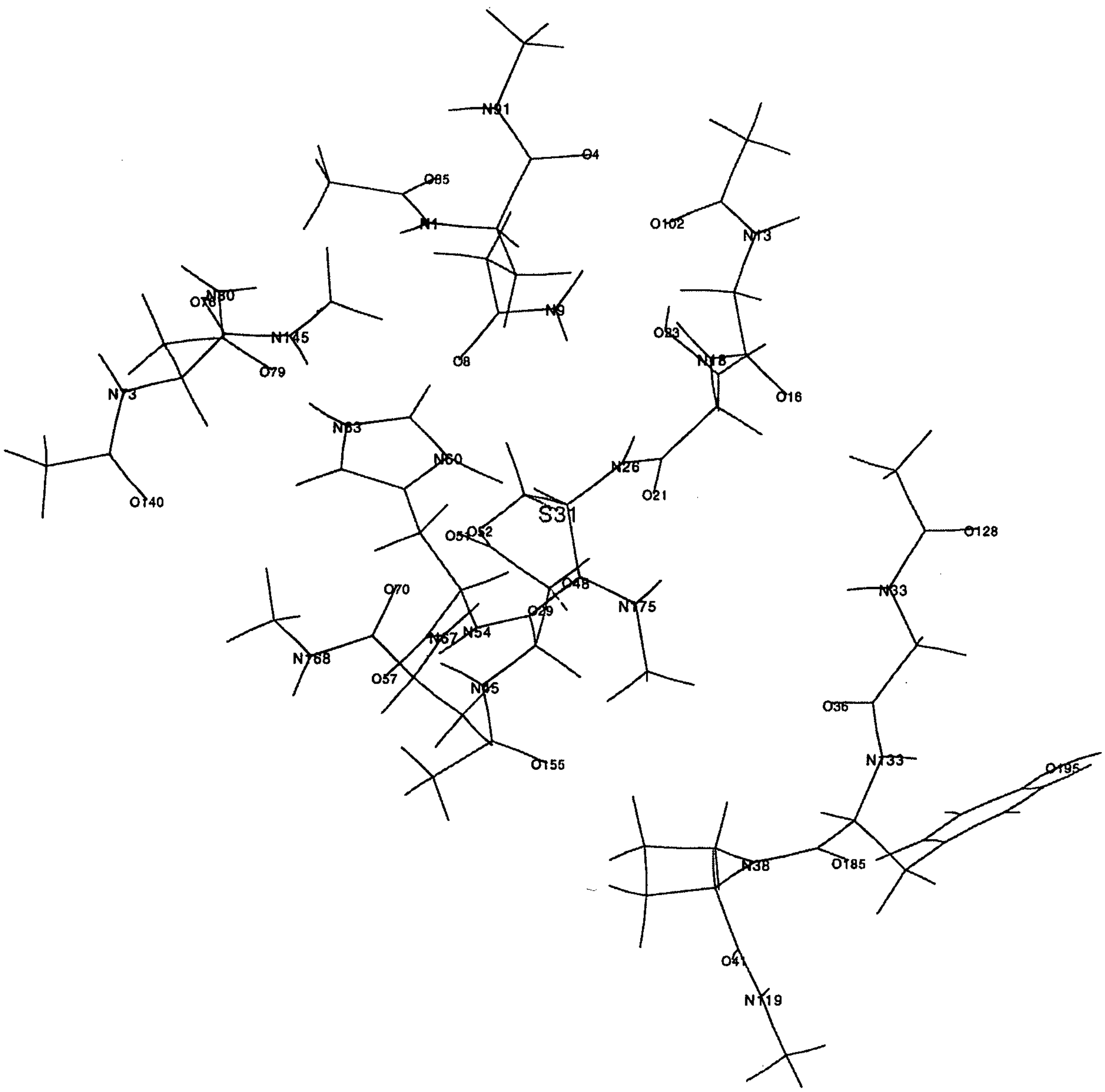


FIGURE 4. A model of the active site of papain obtained by X-ray crystallography (PDB, [8]).

catalytic action of papain. In the R-domain, one can observe a torsion of the plane of the carboxyl group of the lateral chain of Asp158. Since no similar rotation of the terminal bond of the Asp158 fragment is observed, this torsion of the carboxyl group (+30° from the crystal to the active site) could be associated with the mechanism of complex formation. Nevertheless, as pointed out by Dijkman and van Duijnen [30], the effect of the negative charge of Asp158 in the proton-transfer mechanism can be nearly completely screened by the solvent. A second important observation is the torsion of the lateral chain of His159 away from the S—H residue of Cys25, complementary to a similar movement of the —CH_2SH residue.

To optimize the leupeptin–papain complex, we initially moved the sulfur atom from the original crystal position. This leads to reconstruction of the —CHO group of leupeptin. Subsequent AM1 optimization (Fig. 6) moved the sulfur atom back to a

TABLE II
Summary of the relevant dihedral angles of the amino acid fragments used in the model.

	Model I[a]		Model II[b]	
	Crystal structure	AM1-optimized	Crystal structure	AM1-optimized
Gln19				
N9—C7—C6—C5	89.7	66.7	—	66.4
C7—C6—C5—C2	170.0	148.7	—	164.6
C6—C5—C2—N1	−57.1	−71.2	—	−63.6
C5—C2—N1—C84	−166.1	−138.2	—	−148.7
C5—C2—C3—N91	−104.6	−107.2	—	178.0
Gly23				
N18—C15—C14—N13	−66.9	−53.2	—	−55.8
C15—C14—N13—C101	78.1	85.0	—	83.7
Ser24				
N26—C20—C19—C22	−138.0	174.2	—	158.5
C20—C19—C22—O23	73.3	67.3	—	71.3
C22—C19—C22—C15	−132.6	−150.4	—	−115.0
H25—O23—C22—C19	174.2	−72.3	—	−69.0
Cys25				
S31—C30—C27—N26	−63.5	−44.1	−80.3	−54.2
C30—C27—N26—C20	−170.9	−83.2	179.9	−129.3
N175—C28—C27—C30	−97.1	−52.7	38.0	−68.0
Gly66				
C127—N33—C34—C35	173.8	128.7	—	128.6
N33—C34—C35—N133	161.6	−163.1	—	−168.0
Tyr67				
C35—N133—C135—C186	−111.0	−139.8	—	−130.2
N133—C135—C184—C185	52.7	62.9	—	57.8
C186—C135—C184—C135	−60.8	−52.9	—	−61.3
C192—C187—C186—C135	79.7	72.5	—	64.4
Pro68				
N119—C40—C39—N38	39.2	−59.9	—	−65.6
C40—C39—N38—C184	58.0	88.8	—	87.0
Asp158				
O52—C50—C49—C46	136.6	−158.6	—	−127.2
C50—C49—C46—N45	−55.0	−54.7	—	−64.2
C49—C46—N45—C154	−93.0	−136.6	—	−134.9
His159				
N60—C59—C58—C55	−57.0	−49.5	—	−13.7
C59—C58—C55—N54	178.0	169.5	—	−168.9
C58—C55—N54—C47	−88.6	−153.4	—	−102.6
C55—N54—C47—C46	166.1	178.8	—	179.6
N54—C47—C46—N45	−.1	54.0	—	−10.6
O48—C47—C46—N45	−177.9	−125.0	—	171.7
Ala160				
C71—C68—N67—C56	−127.2	−156.2	—	−140.6
C68—N67—C56—C55	178.8	179.5	—	−175.7
N67—C56—C55—N54	−136.4	−77.3	—	−128.4
O57—C56—C55—N54	41.0	103.8	—	31.8

TABLE II
Continued.

	Model I[a] Crystal structure	Model I[a] AM1-optimized	Model II[b] Crystal structure	Model II[b] AM1-optimized
Asn175				
N80 — C78 — C77 — C74	162.5	169.9	—	164.1
C78 — C77 — C74 — N73	169.8	158.1	—	162.5
N145 — C75 — C74 — N73	−157.2	−150.4	—	−147.6
C139 — N73 — C74 — C77	−124.1	−75.4	—	−108.9

[a]Model I — Crystal structure excluding leupeptin and solvent molecules (from PDB [8]).
[b]Model II — crystal structure including leupeptin (from PDB [8]).

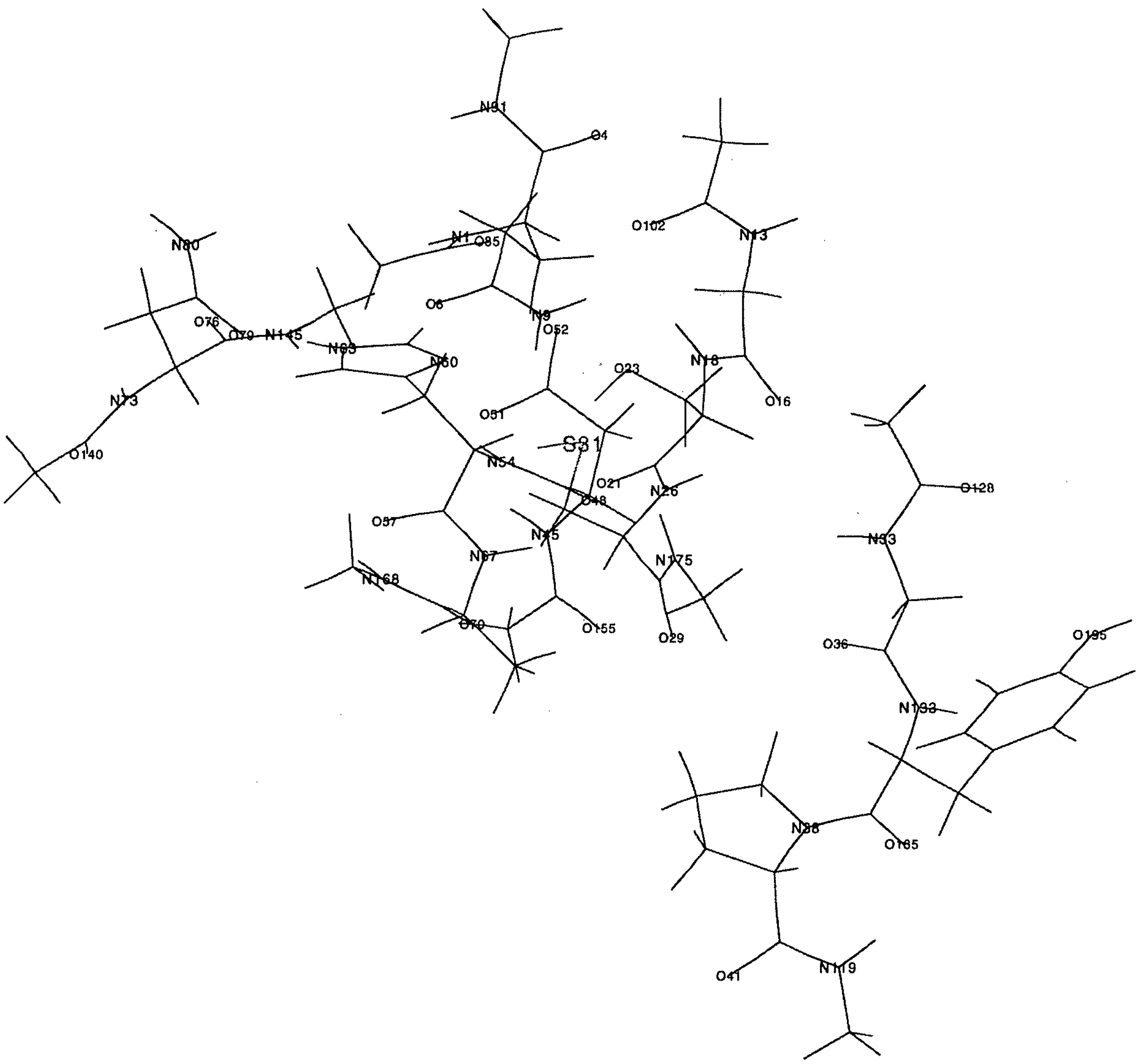


FIGURE 5. An optimized (AM1) model of the active site of papain.

position very close to the original position without proton transfer from the imidazole ring to the sulfur atom. The imidazole ring, however, rotates by +40° (from the crystal structure) in relation to the C—S bond axis. More important, the ImH^+ group is displaced by 0.7 Å, away from the carboxyl group of Asp158, and at the same time, the other N—H—H group of the imidazole ring moves closer (1.2 Å) to Asn175. These results seem to imply that the host molecule is essential for proton transfer from S—H to imidazole. On the other hand, protonation of imidazole increases the strength of the N—H (imidazole) ··· O = C(Asn175) interaction, which is in line with earlier findings [44]. These results, and the fact that proton transfer from imidazole to sulfur occurs in the absence of leupeptin, seem to point either to a concerted mechanism or to proton transfer in a posterior step of complex formation.

Table III shows relevant distances of the complex leupeptin–papain (crystal and AM1-optimized structures) and compares them with a re-

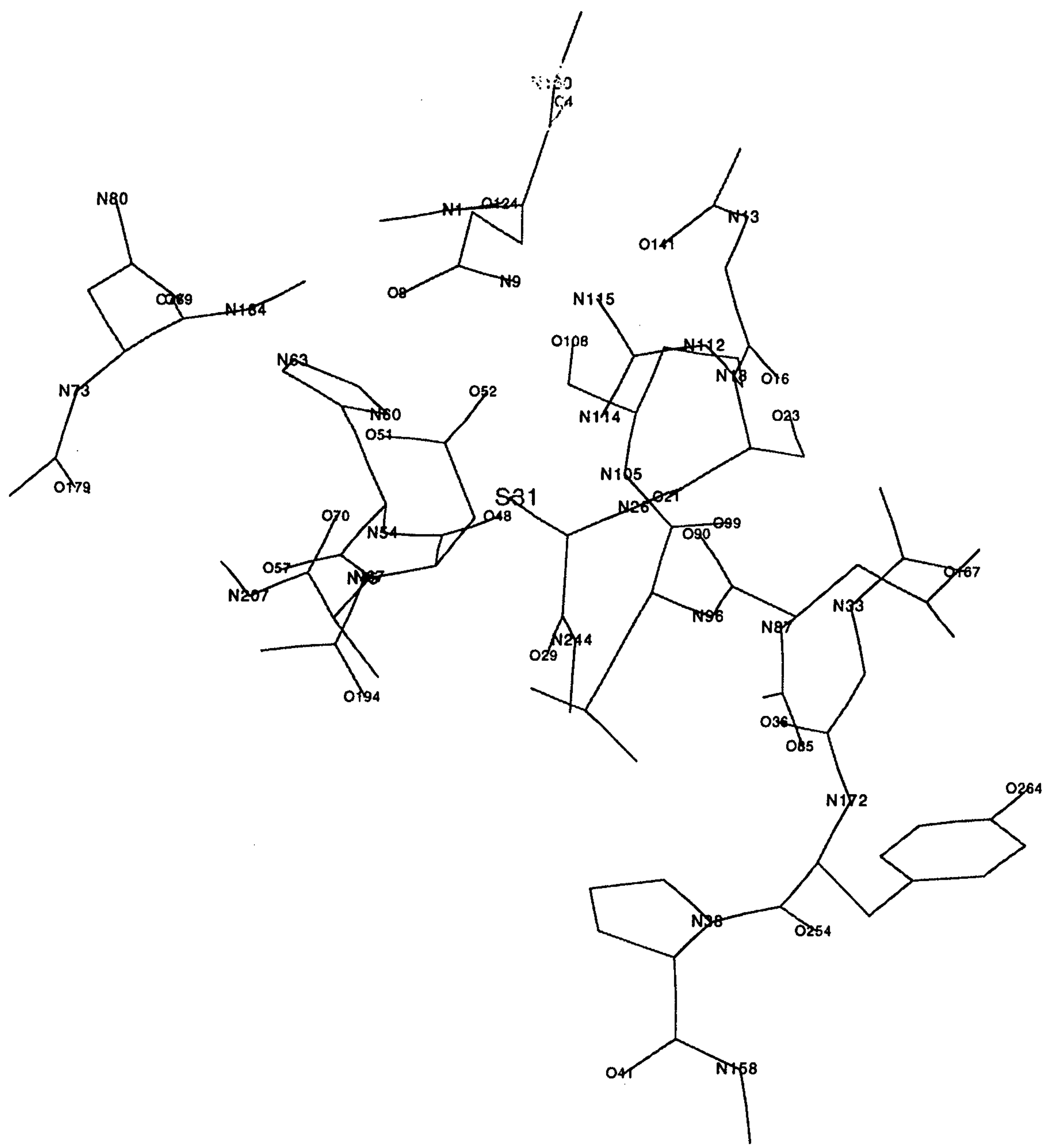


FIGURE 6. An optimized (AM1) model of the active site of the complex papain–leupeptin.

TABLE III
Summary of papain – leupeptin interactions.

Leupeptin residue	Leupeptin atom	Amino acid	Amino acid atom	R, γ crystal[a] A (°)	R, γ AM1 A (°)	[32][b] (A)	[47][c] (A)
P1(Arg)[d]	Oxyanion O	Gln19	Amide N of side chain	2.88[e] (9.7)[e]	2.95 (32.9)	3.10	2.87
P1(Arg)	Oxyanion O	Asp158	Carbonyl O of backbone	3.54 (21.0)	3.17 (18.8)	—	—
P2(Leu)	Carbonyl O of backbone	Gly66	Amide N of backbone	2.78 (14.9)	3.04 (22.0)	—	3.04
P2(Leu)	Amide N of backbone	Gly66	Carbonyl O of backbone	3.05 (10.9)	3.27 (18.9)	—	—

[a]Coordinates obtained from Protein Data Base (DPB) (see [8]).
[b]Correspondent atoms in the calculated active centers (estimated values).
[c]E-64 –papain complex crystal structure.
[d]For nomenclature, see [22].
[e]R and γ are the H-bond parameters as defined by Schuster [48].

cently calculated model [32] and with the crystal structure of an E-64-papain complex [49]. AM1 is known to underestimate hydrogen-bonding formation, consistently giving distances and angles larger than the experimental values [36]. The calculated AM1 values for the leupeptin–papain complex reflects this fact. The exception is the NH $\cdots$ O=C (backbone of leupeptin to backbone of papain interaction). In this case, complex formation increases the hydrogen-bond strength. This can be understood on the grounds of steric interactions.

Concluding Remarks

In this article, we described our modeling of the protease inhibitor leupeptin and its interactions with the proteolytic enzyme papain. The results of our calculations show that the geometry of leupeptin in the crystal is less stable than our best calculated geometry by about 6.0 kcal/mol. Stabilization of the more tensioned structure must come from additional interactions between the host molecule and the peptide backbone lateral chains or solvent molecules in the active center.

Our calculations in the active center of papain show that proton transfer to the thiol group is preferred, according to recent calculations at the SCF/6-31G* level by Beveridge (models 7 and 8 of [33]). Our calculations show also a torsion of the carboxyl group (+30° from the crystal) that could be involved in the mechanism of complex formation. This interaction, however, could be screened by solvent molecules [30]. The calculations also show a displacement of the imidazole ring of His159 away from the S—H residue of Cys25.

The calculations on the leupeptin–papain complex show that reconstruction of the aldehyde–CHO bond of leupeptin led to a leupeptin–papain complex in which the $S^- \cdots HIm^+ \cdots O{=}C$(Asn175) catalytic triad is strengthened. Moreover, comparison of some relevant interactions between leupeptin and papain show that AM1 calculations reproduce fairly well experimental hydrogen-bond distances and angles.

ACKNOWLEDGMENTS

This research received partial financial support from Brazilian agencies CNPq and FAPERJ. We are grateful to the Núcleo de Atendimento em Computação de Alto Desempenho (NACAD-COPPE-UFRJ) for the grant of computational time in their Cray J90. The authors also thank Prof. Michael C. Zerner for the use of the computational facilities at the Quantum Theory Project (Florida).

References

1. A. R. Kennedy, Cancer Res. **54**(7) (Suppl.), 1999s (1994).
2. G. Leto, F. M. Tumminello, N. Gebbia, B. Woynarowska, and R. J. Bernacki, Anticancer Res. **10**(1), 265 (1990).
3. J. Brtko, J. Knopp, and M. E. Baker, Mol. Cell. Endocrinol. **93**, 81 (1993).
4. I. Eto and C. J. Grubbs, Biochem. J. **283**, 209 (1992).
5. I. V. Kurinov and R. W. Harrison, Prot. Sci. **5**, 752 (1996).

6. H. Umezawa, Acta Biol. Med. Ger. **36**(11–12), 1899 (1977).
7. H. Umezawa, Annu. Rev. Microbiol. **36**, 75 (1982).
8. E. Schröder, C. Phillips, E. Garman, K. Harlos, and C. Crawford, FEBS Lett. **315**(1), 38 (1993).
9. D. E. Hudson-Taylor, S. A. Dolan, F. W. Klotz, H. Fujioka, M. Aikawa, E. V. Koonin, and L. H. Miller, Mol. Microbiol. **15**, 463 (1995).
10. J. J. Marr and M. Muller, Eds., *Biochemistry and Molecular Biology of Parasites* (Academic Press, London, 1995).
11. B. C. Elford, G. M. Cowan, and D. J. P. Ferguson, Biochem. J. **308**, 361 (1995).
12. T. E. Wellems, Parasit. Today **7**, 110 (1991).
13. W. H. Wensdorfer and D. Payne, Pharm. Therap. **50**, 95 (1991).
14. I. Ansorge, D. Jeckel, F. Wieland, and K. Lingelbach, Biochem. J. **308**, 335 (1995).
15. J. H. McKerrow, Exp. Parasitol. **68**, 111 (1989).
16. J. D. Lonsdale-Eccles, G. W. N. Mpinbaza, Z. R. M. Nkhungulu, J. Olobo, L. Smith, O. M. Tosomba, and D. J. Grab, Biochem. J. **305**, 549 (1995).
17. J. S. Bond and E. Butler, Annu. Rev. Biochem. **56**, 333 (1987).
18. J. H. McKerrow, E. Sun, P. J. Rosenthal, and J. Bouvier, Annu. Rev. Microbiol. **47**, 821 (1993).
19. C. M. R. de Sant'Anna, R. Bicca de Alencastro, C. R. Rodrigues, G. Barreiro, E. Barreiro, J. D. da Motta Neto, and A. C. C. Freitas, Int. J. Quantum Chem. Quantum Biol. Symp. **23**, 111 (1996).
20. J. Drenth, K. H. Kalk, and H. M. Swen, Biochemistry **15**, 3731 (1976).
21. I. G. Kamphuis, K. H. Kalk, M. B. A. Swarte, and J. Drenth, J. Mol. Biol. **179**, 233 (1984).
22. A. C. Storer and R. Ménard, Methods Enzymol. **244**, 486 (1994).
23. E. N. Baker and J. Drenth, in *Biological Macromolecules and Assemblies*, Vol. 3: *Active Sites of Enzymes*, F. A. Jurnak and A. McPherson Eds. (Wiley, New York, 1987), pp. 313-368.
24. V. Turk and N. Bode, in *Innovations in Proteases and Their Inhibitors*, F. X. Avilés, Ed., (W. de Gruyter, New York, 1993), pp. 161–178.
25. W. Bode and R. Huber, in *Innovations in Proteases and Their Inhibitors*, F. X. Avilés, Ed., (W. de Gruyter, New York, 1993), pp. 81–122.
26. R. Lavéry, A. Pullman, and Y. K. Wen, Int. J. Quantum Chem. **24**(4), 353 (1983).
27. P. Th. van Duijnen, B. Th. Thole, R. Broer, and W. C. Nieuwpoort, Int. J. Quantum Chem. **17**, 651 (1980).
28. J. A. C. Rullmann, M. N. Bellido, and P. Th. van Duijnen, J. Mol. Biol. **206**, 101 (1989).
29. D. Arad, R. Langridge, and P. A. Kollman, J. Am. Chem. Soc. **112**, 491 (1990).
30. J. P. Dijkman and P. Th. van Duijnen, Int. J. Quantum Chem. Quantum Biol. Symp. **18**, 49 (1991).
31. G. D. Duncan, C. P. Huber, and W. J. Welsh, J. Am. Chem. Soc. **114**, 5784 (1992).
32. M. J. Harrison, N. A. Burton, I. H. Hillier, and I. R. Gould, Chem. Commun. 2769 (1996).
33. A. J. Beveridge, Prot. Sci. **5**, 1355 (1996).
34. Y. Lin and W. J. Welsh, J. Mol. Graph. **14**, 62 (1996).
35. N. Swamy Kandadai and M. Rami Reddy, J. Comput. Chem. **17**, 1328 (1996).
36. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, J. Am. Chem. Soc. **107**(13), 3902 (1985).
37. M. J. S. Dewar and E. G. Zoebisch, J. Mol. Struct. (Theochem) **180**, 1 (1988).
38. A. A. Voityuk, J. Struct. Chem. **29**(1), 120 (1988).
39. J. J. P. Stewart, MOPAC 7.0 (Frank J. Seiler Research Lab., US Air Force Academy, Colorado Springs, 1993).
40. A. N. Glazer and E. L. Smith, in *The Enzymes*, 3rd ed., P. D. Boyder, Ed. (Academic Press, New York, 1971), Vol. 3.
41. P. J. Berti, C. H. Faerman, and A. C. Storer, Biochemistry **30**, 1394 (1991).
42. I. G. Kamphuis, K. H. Kalk, M. B. A. Swarte, and J. Drenth, J. Mol. Biol. **179**, 233 (1984).
43. R. M. Schultz, P. Varma-Nelson, R. Ortiz, K. A. Kozlowski, A. T. Orawski, P. Pagast, and A. Frankfater, J. Mol. Biol. **264**(3), 1497 (1989).
44. S. L. Bearne and R. Wolfenden, CHEMTRACTS-Org. Chem. **8**, 288 (1995).
45. R. Ménard, J. Carrière, P. Laflamme, C. Plouffe, H. E. Khouri, T. Vernet, D. C. Tessier, D. Y. Thomas, and A. C. Storer, Biochemistry **30**, 8924 (1991).
46. A. Berger and I. Schechter, Philos. Trans. R. Soc. Lond. B **257**, 249 (1970).
47. E. N. Baker and J. Drenth, in *Biological Macromolecules and Assemblies*, Vol. 3: *Active Sites of Enzymes*, F. A. Jurnak and A. McPherson, Eds. (Wiley, New York, 1987), pp. 313–368.
48. P. Schuster, in *The Hydrogen Bond—Recent Developments in Theory and Experiments*, P. Schuster, G. Zundel, and C. Sàndorfy, Eds. (North-Holland, Amsterdam, 1976), pp. 25–163.
49. K. I. Varughese, F. R. Ahmed, P. R. Carey, S. Hasnain, C. P. Huber, and A. C. Storer, Biochemistry **28**, 1330 (1989).

Theoretical Studies of Inclusion Complexes of α- and β-Cyclodextrin with Benzoic Acid and Phenol

MING-JU HUANG,[1] JOHN D. WATTS,[2] NICHOLAS BODOR[1]

[1] *Center for Drug Discovery, College of Pharmacy, P.O. Box 100497, Health Science Center, Gainesville, Florida 32610-0497*

[2] *Quantum Theory Project, P.O. Box 118435, 362 Williamson Hall, University of Florida, Gainesville, Florida 32611-8435*

Received 2 June 1997; accepted 19 August 1997

ABSTRACT: A series of semiempirical molecular orbital calculations using the AM1 method were performed on the inclusion complexes of α- and β-cyclodextrin with benzoic acid and phenol in the "head-first" and "tail-first" positions. The AM1 results show that α-cyclodextrin complexes with both guest compounds in the "head first" position are more stable than in the "tail-first" position, while the β-cyclodextrin complex with phenol in the "tail-first" position is more stable, but with benzoic acid, the "head-first" position is more stable. The driving forces for complex formation were investigated based on different intramolecular and intermolecular interactions. In addition, 1SCF AM1 calculations were performed on the β-cyclodextrin complexes with benzoic acid in the "tail-first" and "head-first" positions with the benzoic acid moved stepwise along the Z-axis of the β-cyclodextrin principal axis coordinate system. © 1997 John Wiley & Sons, Inc. Int J Quant Chem **65**: 1135–1152, 1997

Introduction

Cyclodextrins (CDs) were first isolated in 1891 by Villiers [1] as degradation products of starch, Schardinger characterized them as cyclic oligosaccharides in 1904 [2], then in 1935, Freudenberg and Jacobi [3] described them as being macrocyclic compounds, built from glucopyranose units linked by α-(1,4)-glycosidic bonds. The most common CDs are α-CD, β-CD, and γ-CD, consisting of six, seven, and eight α-D-glucopyranosyl residues, respectively. The CD molecules have an endolipophilic cavity, which is made water-soluble by many outward-pointing OH groups. This presence of a hydrophobic central cavity enables many different (organic, inorganic, neutral, and ionic) molecules to be incorporated into the cavity, both in the solid state and in solution. Furthermore, complex formation of CDs with guest compounds

Correspondence to: N. Bodor.

© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061135-18

such as drugs and insecticides causes new physicochemical features, leading to practical usages in pharmaceutical chemistry, food technology, and other industrial areas [4, 5].

The considerable power of complexation of the CDs was first recognized by Freudenberg and Cramer [6] in the late 1940s, substantiated by complexation studies by Cramer, Saenger, and others [4, 7]. Evidence that the incorporated guest is sitting at the center of the cavity was obtained by determining the association constant for the solvated complex and confirmed by X-ray crystallography. Since then, many methods have been used to study CDs and their inclusion complexes, including X-ray crystallography [8] for the solid-state crystallographic analyses, NMR spectroscopy [9] for the solution-phase spectroscopic studies which are usually performed in an aqueous environment or involve polar organic solvents, ESR spectroscopy [10], and electrochemical methods [11].

There are no covalent bonds formed or broken during the complex formation process and the complexed molecules are in equilibrium with uncomplexed molecules in aqueous solutions. The driving forces for the complex formation have been attributed to the release of entropy-rich water molecules from the cavity [12], van der Waals interactions [13, 14], hydrogen bonding [15–17], hydrophobic interactions [14, 18], release of ring strain in the cyclodextrin molecule [16], and changes in solvent–surface tension [19]. However, the relative contributions and even the nature of the different forces are not well known. The thermodynamic parameters enthalpy (ΔH) and entropy (ΔS) can be obtained from the temperature dependence of the stability constant of the CD complex. ΔH is always negative but ΔS can be positive or negative [4]. The central cavity of the CD molecule is relatively hydrophobic but CDs are able to form inclusion complexes with a wide variety of compounds ranging from very polar inorganic ions to completely nonpolar organic molecules. The hydrophobicity alone cannot fully explain the complexation. CD complexes are usually studied by NMR spectroscopy, but the kinetics of complex formation introduce some difficulty in the extraction of conclusions from these experiments [20]. We studied the inclusion of a series of phenol and benzoic acid in α-CD and β-CD using the AM1 semiempirical molecular orbital method [21]. We investigated the driving forces of complex formation based on a combination of several intermolecular interactions and possibly hydrophobic effects such as steric fit or size selectivity (this would be the primary criterion), van der Waals interactions, hydrogen bonding, dispersive forces, dipole–dipole interactions, charge-transfer interactions, and electrostatic interactions.

Recently, there have been several theoretical studies of CDs and CD inclusion complexes. A molecular modeling study of structural effects on the binding of amine drugs with the diphenylmethyl functionality to CDs was reported by Tong et al. [22]. A conformational study of β-CD complexes with nootropic drugs using molecular mechanics (MM) was performed by Amato et al. [23]. Conformational analysis of β-CD complexes with a variety of molecules using the MM2 force field was performed by Kostense et al. [24]. Some other molecular mechanical studies of CDs have also been reported [25–35]. The conformational behavior of complexes of α-CD with *p*-chlorophenol and *p*-hydroxybenzoic acid in water was studied using molecular dynamics simulations by van Helden et al. [36]. In addition, several molecular dynamics studies of CDs were reported [37–42]. A series of *fixed-geometry* quantum chemical studies of CDs were performed with the semiempirical CNDO or CNDO/2 methods [43–46]. We recently performed a series of semiempirical calculations on α- and β-CDs using the AM1 method [21], including full geometry optimization [47]. We now report a series of AM1 calculations on complexes with benzoic acid and phenol. In contrast to the CNDO studies, these calculations include complete and unrestricted geometry optimization and conformational analysis. The aim of this study was to compare the stabilization energy and to investigate the driving force for the inclusion of phenol and benzoic acid in α-CD and β-CD.

Methods

AM1 calculations on the inclusion complexes of phenol and benzoic acid with α-CD and β-CD and on the free CDs were performed using a modified version of the AMPAC program from QCPE [21]. Phenol and benzoic acid were studied using the MM2 and MOPAC programs on the Tektronix CAChe (Computer Assisted Chemistry) workstation. MM2 calculations were run on the CAChe workstation to determine starting geometries for the AM1 calculations. The size and conformationally flexibility of the molecules of the present study

make it impossible to study all possible conformers. To find the lowest-energy structures, we tried several starting points for the AM1 optimizations, including some determined by molecular dynamics simulations.

Results and Discussion

Calculated partition coefficients, volume, surface area, ovality, and lengths of the guest compounds, benzoic acid and phenol, are given in Table I. The geometrical data were determined from the AM1-optimized geometries and van der Waals radii. The partition coefficients were calculated by the BLOGP method [48] and depend on geometrical and electronic parameters (such as atomic charges and dipole moment). Table II contains further molecular dimensions of benzoic acid and phenol.

Table III contains enthalpies of formation (ΔH_f), dipole moments, HOMO and LUMO energies of phenol, benzoic acid, α-CD, and β-CD, and inclusion complexes of phenol and benzoic acid with α-CD and β-CD. All data refer to AM1-optimized geometries.* Enthalpies of complexation are also given. These are defined by

$$\Delta\Delta H_f = \Delta H_f(\text{CD} + \text{guest}) - \Delta H_f(\text{CD}) - \Delta H_f(\text{guest}).$$

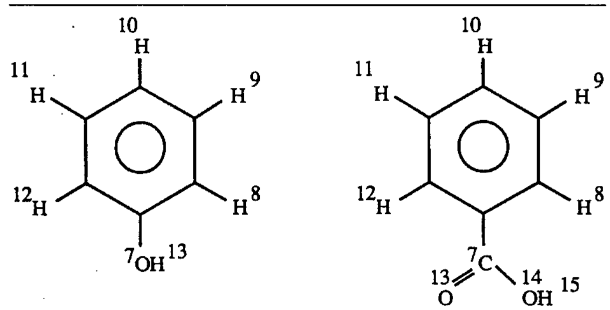
Thus, a negative value for $\Delta\Delta H_f$ indicates that complexation is thermodynamically favorable. The "head-first" and "tail-first" orientations are depicted in Figure 1.

Table III shows that stable inclusion complexes of both guests with both α- and β-CD can be formed. Molecular-size considerations show that formation of inclusion complexes of benzoic acid and phenol is plausible. For example, Van Hooidonk and Breebaart-Hansen [50] estimated the diameter of the β-CD cavity to be 7.5 Å, which is to be compared with the widths of benzoic acid and phenol that may be deduced from the internuclear distances in Table II. Orientation other than "head-first" or "tail-first," i.e., "crosswise" are less likely from size considerations, and no evidence for these was found in the calculations. Experimentally, Siimer et al. [51, 52] showed that both α-CD and β-CD form stable 1 : 1 complexes with benzoic acid. Interestingly, γ-CD forms unstable 1 : 1 and 1 : 2 inclusion complexes.

Looking at the enthalpies of complexation in Table III, we see a number of trends: (1) the hy-

*The ΔH_f values for free α- and β-CD are slightly lower (2 and 9 kcal/mol) than those obtained in our previous study of free cyclodextrins. Apparently, the structures found previously were only local minima and not global minima.

TABLE II
Molecular dimensions of phenol and benzoic acid.



	Phenol	Benzoic acid
R(8–11)	4.997	4.994
R(9–12)	5.003	5.001
R(10–13)	5.662	
R(8–12)	4.336	4.327
R(9–11)	4.322	4.324
R(13–14)		2.210
R(10–15)		7.012

TABLE I
Calculated partition coefficients, volume (Å³), surface area (Å²), ovality, and height (Å) of phenol and benzoic acid.

Compound	Log *P*	Volume	Surface	Ovality	Height
(a) Phenol	1.299	91.84	120.50	1.224	5.662
(b) Benzoic acid	1.610	110.74	143.02	1.282	7.012

TABLE III

Heats of formation (ΔH_f, kcal / mol), dipole moment (Debye), HOMO energy (eV), and LUMO energy (eV) for AM1-optimized geometries of benzoic acid, phenol, α-CD, β-CD, and complexes of α- and β-CD with benzoic acid and phenol in the "head-first" and "tail-first" positions; stabilization energies ($\Delta\Delta H_f$) are in kcal / mol.

Compound	ΔH_f	Dipole	HOMO	LUMO	$\Delta\Delta H_f$
Benzoic acid	−68.0[a]	2.4	−10.08	−0.47	
Phenol	−22.2[a]	1.2	−9.11	0.40	
α-CD	−1416.0	4.0	−10.48	1.58	
β-CD	−1656.6	8.0	−10.34	1.26	
α-CD + benzoic acid "head first"	−1489.1	4.7	−10.21	−0.69	−5.1
α-CD + benzoic acid "tail first"	−1482.0	5.1	−10.26	−0.88	2.0
β-CD + benzoic acid "head first"	−1725.0	6.6	−10.37	−0.82	−0.4
β-CD + benzoic acid "tail first"	−1716.9	8.3	−10.30	−0.83	7.7
α-CD + phenol "head first"	−1439.4	6.6	−9.28	0.26	−1.2
α-CD + phenol "tail first"	−1434.2	8.9	−9.32	0.13	4.0
β-CD + phenol "head first"	−1674.8	2.4	−9.33	0.18	4.0
β-CD + phenol "tail first"	−1679.2	7.5	−9.49	0.09	−0.4

[a]Experimental ΔH_f for phenol and benzoic acid are −23.0 and −70.3 kcal / mol [49].

droxyl group of phenol and the carboxylate group of benzoic acid prefer to face the primary 6-hydroxyl oxygen ("head first") than to face the secondary 2- or 3-hydroxyl oxygen ("tail first") in the α-CD inclusion complexes; (2) the hydroxyl group of phenol prefers to face the secondary 2- or 3-hydroxyl oxygen ("tail first") in the β-CD inclusion complexes, but the carboxylate group of benzoic acid prefers to face the primary 6-hydroxyl oxygen ("head first"); and (3) phenol and benzoic acid form more stable inclusion complexes with α-CD than with β-CD.

For α-CD and benzoic acid, the arrangement in which the carboxylate group of benzoic acid faces the primary 6-hydroxyl oxygen ("head first") has the most stabilization energy (5.1 kcal/mol). This orientational preference is in agreement with previous CNDO or CNDO/2 studies [43–46] for which it was indicated that the dipole moments of guest molecules are antiparallel to the dipole moment of host α-CD in the crystalline state.

Of the four inclusion complexes studied, three prefer the "head-first" orientation and just one, the β-CD and phenol complex, prefers the "tail-first" orientation. The preference of the "tail-first" orientation for the β-CD and phenol complex has also been suggested by others [53–54]. The preference of the "head-first" orientation is in line with

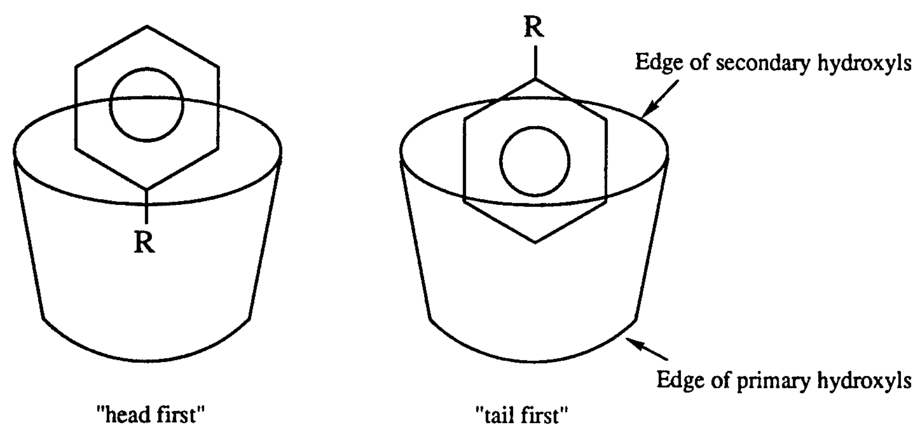


FIGURE 1. The two possible penetration pathways for benzoic acid and phenol.

dipole–dipole interactions. That is not to say, however, that the guest–host interaction can be well modeled by treating the two molecules as point dipoles; specific intermolecular interactions and subtle geometry changes are likely to determine the stability order. One approach is to consider the numbers of intermolecular and intramolecular hydrogen bonds. A hydrogen bond is defined as an O—H ⋯ O interaction in which the H ⋯ O distance is less than or equal to 3.00 Å and the angle at H is larger than 90°. The problems of such cutoff criteria are discussed in greater detail elsewhere [55–58]. The detailed intramolecular hydrogen bonds between the 2-hydroxyl and 3-hydroxyl groups of adjacent glucose units and intramolecular hydrogen bonds between the 2-hydroxyl and 3-hydroxyl groups with their nearby oxygen bridges are listed in Tables IV and V. The numbers of intermolecular hydrogen-bond interactions for the eight different inclusion complexes are shown in Table VI. The total number of intramolecular hydrogen bonds between the 2-hydroxyl and 3-hydroxyl groups of adjacent glucose units (intra 2-3) and the total number of intramolecular hydrogen bonds between the 2-hydroxyl and 3-hydroxyl groups with their nearby oxygen bridges (intra bridge), the total number of intermolecular hydrogen bonds between the host and guest (inter), and the macrocyclic geometric constant O(4) ⋯ O(4′) and O(4) ⋯ O(4′) ⋯ O(4″) are shown in Table VII. From Table VII, we see that the number of intermolecular hydrogen-bond interactions correlates with the orientation of the guest molecule except in the case of the β-CD with benzoic acid. "Head-first" β-CD with benzoic acid has one more intramolecular hydrogen-bond interaction, O9 ⋯ H145—O73, which comes from the 2-hydroxyl group and the oxygen bridge. The "tail-first" orientation has one intermolecular hydrogen-bond interaction, O74—H146 ⋯ O161. However, judging by the distances, the "head-first" orientation probably has stronger hydrogen bonds than does the "tail-first": Compare 2.34 Å for "head first" and 2.46 Å for "tail first."

From Table VII, we see that the macrocyclic constants O(4) ⋯ O(4′) and O(4) ⋯ O(4′) ⋯ O(4″) are slightly increased from α-CD to β-CD, which is in agreement with the results from Saenger [59]. Their values of O(4) ⋯ O(4′) for α- and β-CD are 4.23 and 4.36 Å, respectively, and of O(4) ⋯ O(4′) ⋯ O(4″) for α- and β-CD, 120° and 128°, respectively. As mentioned by Saenger, the macrocyclic geometric constant, the average O(4) ⋯ O(4′) distances (primed atom belongs to the next glucose) forming the edges of the macrocycle, is more or less constant within each member of the CD family, and the individual differences between α- and β-CD arise because the glucose unit has to adjust to the respective radius of the CD.

Fixed geometry AM1 calculations (1SCF) were performed on the host and guest molecules using their geometries in the optimized complex geometry. The enthalpies of formation calculated at these geometries are shown along with their differences from the enthalpies at the optimized geometries in Table VIII. From Table VIII, we find that the most stable inclusion complexes are formed when the host geometry in the complex is closest to the optimized host geometry, except in the case of β-CD with phenol. The 1SCF enthalpy of β-CD at its geometry in the "head-first" β-CD-phenol complex is closer to the enthalpy at the β-CD optimum geometry than is the 1SCF enthalpy at its geometry in the "tail-first" complex, even though the "tail-first" complex has more stabilization energy. We believe this is because that formation of three intermolecular hydrogen bonds in the "tail-first" complex leads to a larger change in the β-CD geometry than in the "head-first" complex.

Comparisons of the AM1 structures of four stable inclusion complexes with those of isolated α- or β-CD are given in Tables IX–XII. Tables IX–XII represent the average values and corresponding ranges of the bond lengths, bond angles, and dihedral angles. The average bond lengths for the four neutral inclusion complexes are not changed from those in the isolated α- or β-CD. The average C2—C3—C4, C3—C4—C5, and O6—C6—C5 bond angles for the α-CD with benzoic acid in the "head-first" position are smaller than those in the isolated α-CD. The average dihedral angles for the α-CD with benzoic acid in the "head-first" position are larger than those in the isolated α-CD, except for the C5—O5—C1—C2 and O5—C1—C2—C3 dihedral angles. The average bond angles and dihedral angles for β-CD with benzoic acid in the "head-first" position are very close to those in the isolated β-CD, except for the C2—C3—C4—C5 and C3—C4—C5—O5 dihedral angles, which are significantly smaller than those in the isolated β-CD. The average bond angles and dihedral angles for the α-CD with phenol with the "head-first" position are very close to those in the isolated α-CD except for the C1—C2—C3, C4—C5—C6,

TABLE IV

Bond distance (Å) for each intramolecular hydrogen-bond interaction between the 2-hydroxyl and the 3-hydroxyl groups of adjacent glucose units of the host molecule in the X-ray conformer, AM1-optimized conformer, and in eight inclusion complexes with the cutoff criteria for hydrogen-bond O — H ··· O interaction with H ··· O distances equal or less than 3.00 Å and angles at H larger than 90°.

Host molecule	No. O—H···O	Atomic no. and H···O bond length	Angles (degrees) at H
α-CD (X-ray)	5	O7 ··· H85 2.18 Å	O7 ··· H85 — O19 = 149.47
		O18 ··· H95 2.03 Å	O18 ··· H95 — O30 = 138.93
		O29 ··· H105 1.99 Å	O29 ··· H105 — O41 = 150.96
		O40 ··· H115 2.00 Å	O40 ··· H115 — O52 = 139.99
		O8 ··· H124 2.08 Å	O8 ··· H124 — O62 = 120.07
α-CD (AM1)	5	O7 ··· H85 2.29 Å	O7 ··· H85 — O19 = 160.70
		O18 ··· H95 2.21 Å	O18 ··· H95 — O30 = 159.80
		O29 ··· H105 2.15 Å	O29 ··· H105 — O41 = 164.31
		O40 ··· H115 2.11 Å	O40 ··· H115 — O52 = 160.99
		O62 ··· H75 2.16 Å	O62 ··· H75 — O8 = 163.55
β-CD (X-ray)	7	O19 ··· H85 2.19 Å	O7 — H85 ··· O19 = 137.60
		O73 ··· H86 2.07 Å	O73 ··· H86 — O8 = 149.50
		O30 ··· H95 2.08 Å	O30 ··· H95 — O18 = 140.70
		O41 ··· H105 2.18 Å	O41 ··· H105 — O29 = 128.28
		O52 ··· H115 2.09 Å	O52 ··· H115 — O40 = 128.59
		O63 ··· H125 2.25 Å	O63 ··· H125 — O51 = 115.26
		(O51 ··· H136 2.53 Å)	(O51 ··· H136 — O63 = 96.46)
		O62 ··· H146 2.03 Å	O62 ··· H146 — O74 = 146.12
β-CD (AM1)	7	O73 ··· H86 2.13 Å	O73 ··· H86 — O8 = 160.45
		O7 ··· H96 2.14 Å	O7 ··· H96 — O19 = 153.22
		O18 ··· H106 2.12 Å	O18 ··· H106 — O30 = 168.43
		O29 ··· H116 2.21 Å	O29 ··· H116 — O41 = 154.27
		O63 ··· H125 2.17 Å	O63 ··· H125 — O51 = 139.26
		O40 ··· H126 2.13 Å	O40 ··· H126 — O52 = 165.37
		O74 ··· H135 2.25 Å	O74 ··· H135 — O62 = 177.47
α-CD + BA "head first"	6	O62 ··· H75 2.13 Å	O62 ··· H75 — O8 = 169.40
		O7 ··· H85 2.24 Å	O7 ··· H85 — O19 = 163.56
		O18 ··· H95 2.27 Å	O18 ··· H95 — O30 = 170.25
		O29 ··· H105 2.20 Å	O29 ··· H105 — O41 = 150.89
		O40 ··· H115 2.17 Å	O40 ··· H115 — O52 = 170.14
		O63 ··· H114 2.90 Å	O63 ··· H114 — O51 = 134.07
		(O51 ··· H125 2.88 Å)	(O51 ··· H125 — O63 = 135.20)
α-CD + BA "tail first"	6	O62 ··· H75 2.10 Å	O62 ··· H75 — O8 = 154.66
		O7 ··· H85 2.12 Å	O7 ··· H85 — O19 = 159.77
		O18 ··· H95 2.08 Å	O18 ··· H95 — O30 = 158.20
		O29 ··· H105 2.10 Å	O29 ··· H105 — O41 = 164.33
		O63 ··· H114 2.14 Å	O63 ··· H114 — O51 = 150.72
		O40 ··· H115 2.12 Å	O40 ··· H115 — O52 = 160.55

TABLE IV
(Continued)

Host molecule	No. O—H···O	Atomic no. and H···O bond length	Angles (degrees) at H
β-CD + BA "head first"	7	O73···H86 2.20 Å	O73···H86—O8 = 153.26
		O7···H96 2.14 Å	O7···H96—O19 = 155.88
		O18···H106 2.12 Å	O18···H106—O30 = 169.59
		O29···H116 2.30 Å	O29···H116—O41 = 153.49
		O63···H125 2.22 Å	O63···H125—O51 = 144.49
		O40···H126 2.21 Å	O40···H126—O52 = 160.52
		O74···H135 2.16 Å	O74···H135—O62 = 167.06
β-CD + BA "tail first"	7	O73···H86 2.17 Å	O73···H86—O8 = 159.37
		O7···H96 2.18 Å	O7···H96—O19 = 150.04
		O18···H106 2.12 Å	O18···H106—O30 = 167.65
		O29···H116 2.22 Å	O29···H116—O41 = 144.78
		O63···H125 2.19 Å	O63···H125—O51 = 140.10
		O40···H126 2.13 Å	O40···H126—O52 = 166.80
		O74···H135 2.21 Å	O74···H135—O62 = 178.60
α-CD + phenol "head first"	6	O19···H74 2.61 Å	O19···H74—O7 = 96.01
		(O7···H85 2.20 Å)	(O7···H85—O19 = 126.00)
		O18···H95 2.16 Å	O18···H95—O30 = 164.62
		O29···H105 2.18 Å	O29···H105—O41 = 158.94
		O40···H115 2.15 Å	O40···H115—O52 = 162.51
		O63···H114 2.31 Å	O63···H114—O51 = 141.24
		O8···H124 2.24 Å	O8···H124—O62 = 145.72
α-CD + phenol "tail first"	6	O7···H85 2.28 Å	O7···H85—O19 = 146.97
		O18···H95 2.11 Å	O18···H95—O30 = 146.78
		O29···H105 2.15 Å	O29···H105—O41 = 166.68
		O40···H115 2.20 Å	O40···H115—O52 = 159.94
		O63···H114 2.14 Å	O63···H114—O51 = 149.78
		O8···H124 2.11 Å	O8···H124—O62 = 156.43
β-CD + phenol "head first"	7	O73···H86 2.13 Å	O73···H86—O8 = 169.94
		O7···H96 2.25 Å	O7···H96—O19 = 151.88
		O18···H106 2.15 Å	O18···H106—O30 = 169.41
		O29···H116 2.15 Å	O29···H116—O41 = 159.42
		O40···H126 2.21 Å	O40···H126—O52 = 167.12
		O51···H136 2.92 Å	O51···H136—O63 = 106.00
		O62···H146 2.11 Å	O62···H146—O74 = 159.49
β-CD + phenol "tail first"	7	O73···H86 2.15 Å	O73···H86—O8 = 149.11
		O7···H96 2.14 Å	O7···H96—O19 = 159.12
		O18···H106 2.14 Å	O18···H106—O30 = 166.70
		O29···H116 2.18 Å	O29···H116—O41 = 144.54
		O63···H125 2.20 Å	O63···H125—O51 = 147.82
		O40···H126 2.19 Å	O40···H126—O52 = 159.69
		O74···H135 2.11 Å	O74···H135—O62 = 175.94

TABLE V

Bond distance (Å) for each intramolecular hydrogen-bond interaction between the 2-hydroxyl or the 3-hydroxyl groups and nearby oxygen bridges of the host molecule in the X-ray conformer, AM1-optimized conformer, and in eight inclusion complexes with the cutoff criteria for hydrogen-bond O — H ⋯ O interaction with H ⋯ O distances equal or less than 3.00 Å and angles at H larger than 90°.

Host molecule	No. O — H ⋯ O	Atomic no. and H ⋯ O bond length	Angles (degrees) at H
α-CD (X-ray)	5	O20 ⋯ H85 2.40 Å	O20 ⋯ H85 — O19 = 105.78
		O31 ⋯ H95 2.44 Å	O31 ⋯ H95 — O30 = 102.18
		O42 ⋯ H105 2.58 Å	O42 ⋯ H105 — O41 = 98.29
		O64 ⋯ H114 2.46 Å	O64 ⋯ H114 — O51 = 102.17
		O53 ⋯ H115 2.66 Å	O53 ⋯ H115 — O52 = 93.07
α-CD (AM1)	7	O9 ⋯ H75 2.73 Å	O9 ⋯ H75 — O8 = 94.06
		O20 ⋯ H85 2.49 Å	O20 ⋯ H85 — O19 = 106.01
		O31 ⋯ H95 2.47 Å	O31 ⋯ H95 — O30 = 105.68
		O42 ⋯ H105 2.73 Å	O42 ⋯ H105 — O41 = 95.11
		O64 ⋯ H114 2.50 Å	O64 ⋯ H114 — O51 = 95.88
		O53 ⋯ H115 2.52 Å	O53 ⋯ H115 — O52 = 103.51
		O64 ⋯ H125 2.46 Å	O64 ⋯ H125 — O63 = 105.04
β-CD (X-ray)	7	O20 ⋯ H85 2.31 Å	O20 ⋯ H85 — O7 = 105.74
		O9 ⋯ H86 2.51 Å	O9 ⋯ H86 — O8 = 97.67
		O31 ⋯ H95 2.28 Å	O31 ⋯ H95 — O18 = 105.14
		O42 ⋯ H105 2.32 Å	O42 ⋯ H105 — O29 = 108.35
		O53 ⋯ H115 2.54 Å	O53 ⋯ H115 — O40 = 95.45
		O64 ⋯ H125 2.51 Å	O64 ⋯ H125 — O51 = 94.95
		O75 ⋯ H146 2.43 Å	O75 ⋯ H146 — O74 = 103.57
β-CD (AM1)	7	O9 ⋯ H86 2.78 Å	O9 ⋯ H86 — O8 = 92.47
		O20 ⋯ H96 2.46 Å	O20 ⋯ H96 — O19 = 107.34
		O31 ⋯ H106 2.64 Å	O31 ⋯ H106 — O30 = 98.64
		O42 ⋯ H116 2.43 Å	O42 ⋯ H116 — O41 = 108.05
		O64 ⋯ H125 2.48 Å	O64 ⋯ H125 — O51 = 101.45
		O53 ⋯ H126 2.63 Å	O53 ⋯ H126 — O52 = 96.92
		O75 ⋯ H135 2.46 Å	O75 ⋯ H135 — O62 = 106.66
α-CD + BA "head first"	7	O9 ⋯ H75 2.62 Å	O9 ⋯ H75 — O8 = 99.86
		O20 ⋯ H85 2.51 Å	O20 ⋯ H85 — O19 = 105.08
		O31 ⋯ H95 2.54 Å	O31 ⋯ H95 — O30 = 102.50
		O42 ⋯ H105 2.44 Å	O42 ⋯ H105 — O41 = 107.44
		O64 ⋯ H114 2.69 Å	O64 ⋯ H114 — O51 = 92.54
		O53 ⋯ H115 2.63 Å	O53 ⋯ H115 — O52 = 99.96
		O64 ⋯ H125 2.44 Å	O64 ⋯ H125 — O63 = 103.55
α-CD + BA "tail first"	7	O9 ⋯ H75 2.82 Å	O9 ⋯ H75 — O8 = 93.93
		O20 ⋯ H85 2.49 Å	O20 ⋯ H85 — O19 = 106.36
		O31 ⋯ H95 2.60 Å	O31 ⋯ H95 — O30 = 102.26
		O42 ⋯ H105 2.58 Å	O42 ⋯ H105 — O41 = 101.76
		O64 ⋯ H114 2.63 Å	O64 ⋯ H114 — O51 = 98.03
		O53 ⋯ H115 2.58 Å	O53 ⋯ H115 — O52 = 102.11
		O9 ⋯ H124 2.51 Å	O9 ⋯ H124 — O62 = 95.57

TABLE V
(Continued).

Host molecule	No. O—H···O	Atomic no. and H···O bond length	Angles (degrees) at H
β-CD + BA "head first"	8	O9···H86 2.80 Å	O9···H86—O8 = 92.12
		O20···H96 2.42 Å	O20···H96—O19 = 108.48
		O31···H106 2.66 Å	O31···H106—O30 = 98.90
		O42···H116 2.43 Å	O42···H116—O41 = 107.90
		O64···H125 2.43 Å	O64···H125—O51 = 104.23
		O53···H126 2.71 Å	O53···H126—O52 = 93.01
		O75···H135 2.46 Å	O75···H135—O62 = 105.14
		O9···H145 2.34 Å	O9···H145—O73 = 101.83
β-CD + BA "tail first"	7	O9···H86 2.80 Å	O9···H86—O8 = 92.47
		O20···H96 2.44 Å	O20···H96—O19 = 108.02
		O31···H106 2.65 Å	O31···H106—O30 = 98.12
		O42···H116 2.45 Å	O42···H116—O41 = 107.31
		O64···H125 2.46 Å	O64···H125—O51 = 102.86
		O53···H126 2.65 Å	O53···H126—O52 = 96.82
		O75···H135 2.46 Å	O75···H135—O62 = 106.00
α-CD + phenol "head first"	6	O20···H85 2.52 Å	O20···H85—O19 = 99.58
		O31···H95 2.76 Å	O31···H95—O30 = 97.56
		O42···H105 2.46 Å	O42···H105—O41 = 106.14
		O64···H114 2.38 Å	O64···H114—O51 = 105.27
		O53···H115 2.56 Å	O53···H115—O52 = 100.58
		O9···H124 2.43 Å	O9···H124—O62 = 108.05
α-CD + phenol "tail first"	5	O20···H85 2.55 Å	O20···H85—O19 = 100.51
		O31···H95 2.52 Å	O31···H95—O30 = 104.74
		O42···H105 2.67 Å	O42···H105—O41 = 97.43
		O64···H114 2.59 Å	O64···H114—O51 = 99.14
		O53···H115 2.52 Å	O53···H115—O52 = 104.22
β-CD + phenol "head first"	8	O9···H86 2.68 Å	O9···H86—O8 = 99.54
		O20···H96 2.43 Å	O20···H96—O19 = 107.71
		O31···H106 2.62 Å	O31···H106—O30 = 99.09
		O42···H116 2.49 Å	O42···H116—O41 = 105.88
		O53···H126 2.69 Å	O53···H126—O52 = 99.28
		O75···H135 2.54 Å	O75···H135—O62 = 90.13
		O64···H136 2.39 Å	O64···H136—O63 = 107.01
		O75···H146 2.73 Å	O75···H146—O74 = 94.47
β-CD + phenol "tail first"	6	O20···H96 2.46 Å	O20···H96—O19 = 107.34
		O31···H106 2.68 Å	O31···H106—O30 = 97.11
		O42···H116 2.46 Å	O42···H116—O41 = 107.81
		O64···H125 2.42 Å	O64···H125—O51 = 104.65
		O53···H126 2.74 Å	O53···H126—O52 = 91.97
		O75···H135 2.53 Å	O75···H135—O62 = 103.24

TABLE VI

Bond distance (Å) for each hydrogen-bond interaction between the guest and host of eight inclusion complexes with the cutoff criteria for hydrogen-bond O — H ··· O interaction with H ··· O distances equal or less than 3.00 Å and angles at H larger than 90°.

Inclusion complexes	No. O — H ··· O	Atomic no. and H ··· O bond length	Angles (degrees) at H
α-CD + BA "head first"	4	O11 ··· H141 (1°) 2.43 Å	O11 ··· H141 — O140 = 119.58
		O33 ··· H141 (1°) 2.37 Å	O33 ··· H141 — O140 = 107.04
		H76 ··· O139 (1°) 2.10 Å	O11 — H76 ··· O139 = 137.80
		H116 ··· O139 (1°) 2.21 Å	O55 — H116 ··· O139 = 158.94
α-CD + BA "tail first"	1	H124 ··· O139 (2°) 2.22 Å	O62 — H124 ··· O139 = 158.24
β-CD + BA "head first"	0		
β-CD + BA "tail first"	1	H146 ··· O161 (2°) 2.46 Å	O74 — H146 ··· O161 = 152.06
α-CD + phenol "head first"	2	O55 ··· H139 (1°) 2.13 Å	O55 ··· H139 — O133 = 124.77
		H76 ··· O133 (1°) 2.14 Å	O11 — H76 ··· O133 = 165.70
		H116 ··· O133 (1°) 2.88 Å	O55 — H116 ··· O133 = 75.62
α-CD + phenol "tail first"	0		
β-CD + phenol "head first"	0		
β-CD + phenol "tail first"	3	H145 ··· O154 (2°) 2.48 Å	O73 — H145 ··· O154 = 96.31
		O73 ··· H160 (2°) 2.26 Å	O73 ··· H160 — O154 = 110.79
		H146 ··· O154 (2°) 2.25 Å	O74 — H146 ··· O154 = 123.19

TABLE VII

Total number of intramolecular hydrogen bonds between the 2-hydroxyl and 3-hydroxyl groups of adjacent glucose units (intra 2 – 3) and total number of intramolecular hydrogen bonds between the 2-hydroxyl and the 3-hydroxyl groups with their nearby oxygen bridges (intra bridge), the total number of intermolecular hydrogen bonds between the host and guest (inter), and the macrocyclic geometric constant O(4) ··· O(4′) and O(4) ··· O(4′) ··· O(4″).

Inclusion complexes	No. intra 2 – 3	No. intra bridges	No. inter	O(4) ··· O(4′)	O(4) ··· O(4′) ··· O(4″)
α-CD (X-ray)	5	5		4.246	119.658
α-CD (AM1)	5	7		4.186	118.781
β-CD (X-ray)	7	7		4.385	128.330
β-CD (AM1)	7	7		4.285	128.301
α-CD + BA "head first"	6	7	4	4.218	119.954
α-CD + BA "tail first"	6	7	1	4.244	119.871
β-CD + BA "head first"	7	8	0	4.276	128.226
β-CD + BA "tail first"	7	7	1	4.286	128.353
α-CD + phenol "head first"	6	6	2	4.223	119.983
α-CD + phenol "tail first"	6	5	0	4.170	119.842
β-CD + phenol "head first"	7	8	0	4.265	127.605
β-CD + phenol "tail first"	7	6	3	4.254	128.325

TABLE VIII

Heats of formation (ΔH_f, kcal / mol) for 1SCF AM1 calculations for the β-CD, benzoic acid, and phenol by removing the host or guest molecules from their optimized configuration; energies differences from their optimized minimum energies ($\Delta\Delta H_f^d$) are in kcal / mol.

Compound	ΔH_f 1SCF for host	$\Delta\Delta H_f^{d1\,a}$	ΔH_f 1SCF for guest	$\Delta\Delta H_f^{d2\,b}$
Benzoic acid (BA)	−68.0			
Phenol	−22.2			
α-CD	−1416.0			
β-CD	−1656.6			
α-CD + BA "head first"	−1411.8	4.2	−67.7	0.3
α-CD + BA "tail first"	−1410.1	5.9	−67.6	0.4
β-CD + BA "head first"	−1655.0	1.6	−67.9	0.1
β-CD + BA "tail first"	−1647.6	9.0	−67.8	0.2
α-CD + phenol "head first"	−1413.2	2.8	−21.8	0.4
α-CD + phenol "tail first"	−1408.9	7.1	−22.1	0.1
β-CD + phenol "head first"	−1652.4	4.2	−22.2	0.0
β-CD + phenol "tail first"	−1650.9	5.7	−22.2	0.0

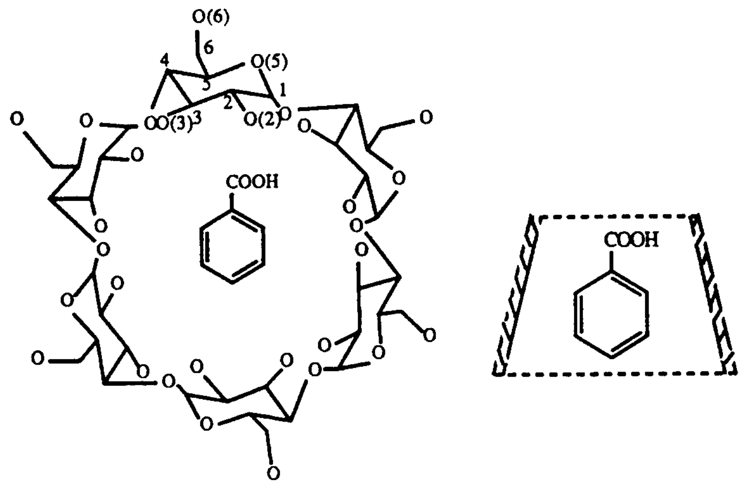
[a] $\Delta\Delta H_f^{d1} = \Delta H_f$ (1SCF for host) − ΔH_f (optimized host).
[b] $\Delta\Delta H_f^{d2} = \Delta H_f$ (1SCF for guest) − ΔH_f (optimized guest).

and O2—C2—C3 bond angles and C1—C2—C3—C4, C5—O5—C1—C2, and O5—C1—C2—C3 dihedral angles. The average bond angles and dihedral angles for the β-CD with phenol in the "tail-first" position are very close to those in the isolated β-CD except for the C1—C2—C3, O5—C1—C2, and O3—C3—C4 bond angles and C1—C2—C3—C4, C5—O5—C1—C2, O5—C1—C2—C3, and O2—C2—C3—C4 dihedral angles.

The final stage of this work involved the calculation of the enthalpy (by single-point AM1 calculation) of the complexes as the guest is moved out of the host. The geometries of guest and host were fixed at their optimum values for the complex and the guest was moved along the host's principal inertial axis which has the largest moment of inertia. This axis is roughly perpendicular to the plane of the CD. We chose it to be the Z-axis of the host principal axis system. Once the geometry of the system is expressed in terms of the host principal axis system, it is trivial to translate the guest along the Z-axis. We define $Z = 0.00$ to be the AM1-optimized geometry. The heats of formation from single-point AM1 calculations for the β-CD with benzoic acid in the "head-first" and "tail-first" positions with the guest moved along Z-axis of the host principal axis coordinate system stepwise are listed in Tables XIII and XIV. In Figures 2 and 3, we show the single-point AM1-calculated energy profile vs. the displacement of the guest along the Z-axis of the host principal axis coordinate system. Figure 2 shows the effect of displacing the benzoic acid from its optimum position in the β-CD with benzoic acid (bottom-first) complex. Movement in the $+Z$ direction corresponds to movement toward the secondary alcohol, and movement in the $-Z$ direction means movement toward the primary alcohol. When the guest moves in the $-Z$ direction, ΔH_f increases more slowly than when the guest moves in the $+Z$ direction. When the guest has moved 3 Å in the $+Z$ direction, the benzene ring is barely interacting with the secondary alcohol, and the rest of the guest is outside of the cavity. Then, when the guest is moved 1 and 2 Å in the $+Z$ direction, the benzene ring is deeper inside the cavity, and the inclusion complex is more stable, until the minimum conformer is reached. After the minimum configuration is reached, moving the guest in the $-Z$ direction increases the heats of formation more slowly than when moving the benzene ring into the cavity. In Figure 3, the behavior of the β-CD with benzoic acid (head-first) complex is shown. When the carboxylic acid group starts to enter the host from the secondary alcohol (+3 Å position), the ΔH_f is about 0.7 kcal/mol higher than the most stable configuration. Then, moving the carboxylic acid group deeper into the host cavity, which means

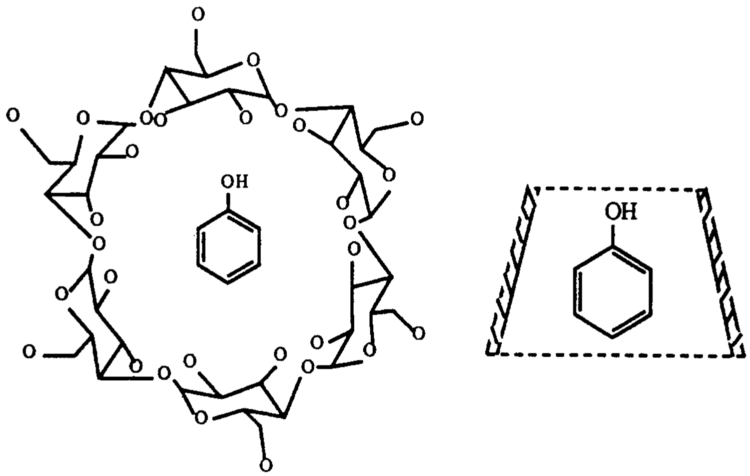
TABLE IX

AM1 structural parameters of C_6O_5 subunits of the inclusion complex of α-CD and benzoic acid and those of α-CD; average values of each parameter and their ranges are given:



	α-CD + benzoic acid		α-CD	
	Ave.	Range	Ave.	Range
Bond length (Å)				
C1—C2	1.543	[1.542, 1.546]	1.542	[1.540, 1.544]
C2—C3	1.534	[1.531, 1.535]	1.533	[1.530, 1.536]
C3—C4	1.540	[1.538, 1.542]	1.539	[1.537, 1.542]
C4—C5	1.537	[1.534, 1.539]	1.536	[1.534, 1.539]
C5—C6	1.534	[1.532, 1.536]	1.534	[1.532, 1.538]
C5—O5	1.432	[1.428, 1.436]	1.432	[1.430, 1.434]
O5—C1	1.411	[1.409, 1.414]	1.412	[1.408, 1.414]
C1—O1	1.415	[1.412, 1.420]	1.414	[1.413, 1.416]
C2—O2	1.414	[1.408, 1.417]	1.414	[1.410, 1.416]
C3—O3	1.416	[1.414, 1.418]	1.417	[1.414, 1.419]
Bond angle (deg.)				
C1—C2—C3	109.6	[108.7, 110.2]	109.6	[109.0, 110.6]
C2—C3—C4	110.2	[109.0, 111.1]	110.5	[109.8, 111.2]
C3—C4—C5	110.8	[109.7, 112.3]	111.3	[109.9, 112.5]
C4—C5—C6	111.6	[110.7, 112.4]	111.7	[110.9, 112.8]
O5—C1—C2	110.8	[110.3, 111.1]	110.9	[110.1, 112.4]
O6—C6—C5	111.4	[110.4, 112.5]	111.5	[110.3, 112.1]
O2—C2—C3	111.6	[110.9, 113.1]	111.5	[110.9, 112.0]
O3—C3—C4	111.0	[110.3, 111.8]	110.9	[110.4, 111.1]
Dihedral angle (deg.)				
C1—C2—C3—C4	−54.3	[−57.2, −51.8]	−53.3	[−56.2, −48.8]
C2—C3—C4—C5	52.1	[47.8, 55.8]	50.9	[48.4, 53.5]
C3—C4—C5—O5	−51.1	[−53.6, −46.8]	−50.3	[−56.1, −44.0]
C4—C5—O5—C1	55.8	[54.1, 57.8]	55.5	[50.2, 60.2]
C5—O5—C1—C2	−58.4	[−60.6, −56.2]	−58.7	[−61.1, −57.8]
O5—C1—C2—C3	56.4	[54.3, 59.4]	56.4	[50.4, 59.5]
O2—C2—C3—C4	−173.7	[−177.4, −170.5]	−172.6	[−175.5, −170.3]
O3—C3—C4—C5	170.2	[166.0, 173.7]	168.7	[166.0, 171.1]

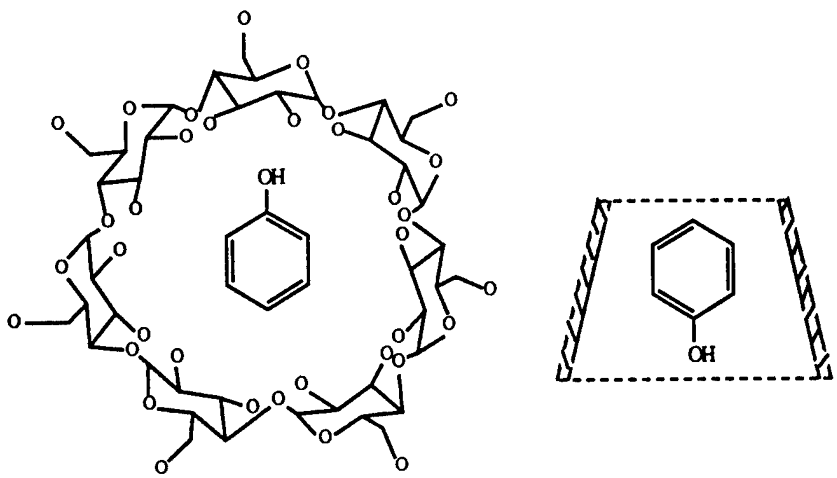
TABLE X
AM1 structural parameters of C_6O_5 subunits of the inclusion complex of α-CD and phenol and those of α-CD; average values of each parameter and their ranges are given:



	α-CD + phenol		α-CD	
	Ave.	Range	Ave.	Range
Bond length (Å)				
C1—C2	1.542	[1.539, 1.544]	1.542	[1.540, 1.544]
C2—C3	1.535	[1.533, 1.537]	1.533	[1.530, 1.536]
C3—C4	1.540	[1.538, 1.543]	1.539	[1.537, 1.542]
C4—C5	1.538	[1.534, 1.539]	1.536	[1.534, 1.539]
C5—C6	1.533	[1.531, 1.535]	1.534	[1.532, 1.538]
C5—O5	1.432	[1.429, 1.436]	1.432	[1.430, 1.434]
O5—C1	1.411	[1.407, 1.415]	1.412	[1.408, 1.414]
C1—O1	1.415	[1.412, 1.418]	1.414	[1.413, 1.416]
C2—O2	1.414	[1.407, 1.421]	1.414	[1.410, 1.416]
C3—O3	1.418	[1.415, 1.419]	1.417	[1.414, 1.419]
Bond angle (deg.)				
C1—C2—C3	110.3	[109.5, 111.8]	109.6	[109.0, 110.6]
C2—C3—C4	110.7	[108.6, 112.2]	110.5	[109.8, 111.2]
C3—C4—C5	111.1	[109.4, 113.4]	111.3	[109.9, 112.5]
C4—C5—C6	110.9	[110.2, 111.8]	111.7	[110.9, 112.8]
O5—C1—C2	111.2	[109.1, 113.0]	110.9	[110.1, 112.4]
O6—C6—C5	111.5	[110.7, 112.1]	111.5	[110.3, 112.1]
O2—C2—C3	110.3	[105.5, 112.1]	111.5	[110.9, 112.0]
O3—C3—C4	109.2	[106.2, 111.6]	110.9	[110.4, 111.1]
Dihedral angle (deg.)				
C1—C2—C3—C4	−52.2	[−57.2, −47.7]	−53.3	[−56.2, −48.8]
C2—C3—C4—C5	50.1	[41.1, 57.5]	50.9	[48.4, 53.5]
C3—C4—C5—O5	−50.0	[−55.9, −40.1]	−50.3	[−56.1, −44.0]
C4—C5—O5—C1	55.0	[50.6, 57.9]	55.5	[50.2, 60.2]
C5—O5—C1—C2	−57.3	[−62.3, −55.0]	−58.7	[−61.1, −57.8]
O5—C1—C2—C3	54.9	[50.7, 60.4]	56.4	[50.4, 59.5]
O2—C2—C3—C4	−173.2	[−176.3, −169.3]	−172.6	[−175.5, −170.3]
O3—C3—C4—C5	168.8	[159.1, 174.7]	168.7	[166.0, 171.1]

TABLE XI

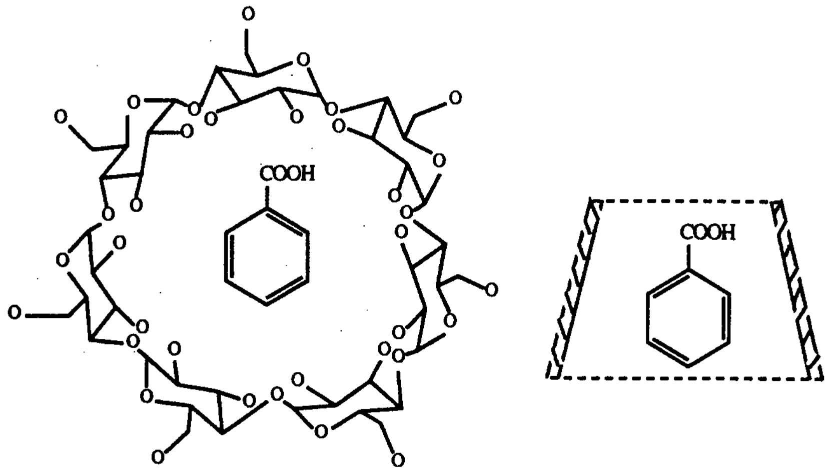
AM1 structural parameters of C_6O_5 subunits of the inclusion complex of β-CD and phenol and those of β-CD; average values of each parameter and their ranges are given:



	β-CD + phenol		β-CD	
	Ave.	Range	Ave.	Range
Bond length (Å)				
C1—C2	1.541	[1.538, 1.545]	1.541	[1.538, 1.544]
C2—C3	1.536	[1.534, 1.537]	1.535	[1.534, 1.536]
C3—C4	1.538	[1.537, 1.540]	1.538	[1.537, 1.540]
C4—C5	1.538	[1.535, 1.541]	1.537	[1.535, 1.538]
C5—C6	1.533	[1.531, 1.536]	1.533	[1.532, 1.534]
C5—O5	1.432	[1.430, 1.433]	1.431	[1.430, 1.432]
O5—C1	1.412	[1.407, 1.417]	1.411	[1.405, 1.414]
C1—O1	1.418	[1.413, 1.425]	1.417	[1.413, 1.421]
C2—O2	1.413	[1.407, 1.417]	1.413	[1.406, 1.416]
C3—O3	1.417	[1.416, 1.418]	1.417	[1.416, 1.418]
Bond angle (deg.)				
C1—C2—C3	109.8	[109.0, 110.4]	110.4	[109.6, 112.1]
C2—C3—C4	109.7	[108.4, 110.4]	110.0	[108.9, 111.2]
C3—C4—C5	110.5	[108.4, 112.0]	110.6	[110.1, 111.6]
C4—C5—C6	111.6	[110.3, 113.8]	111.3	[110.8, 111.8]
O5—C1—C2	111.4	[110.6, 112.8]	112.0	[111.0, 113.8]
O6—C6—C5	112.3	[111.7, 114.3]	112.0	[111.6, 112.5]
O2—C2—C3	111.2	[110.4, 112.1]	111.2	[110.3, 112.4]
O3—C3—C4	110.5	[107.1, 112.8]	109.9	[106.8, 112.0]
Dihedral angle (deg.)				
C1—C2—C3—C4	−54.5	[−57.9, −51.6]	−52.7	[−54.5, −46.7]
C2—C3—C4—C5	53.6	[49.6, 58.7]	53.3	[51.8, 55.6]
C3—C4—C5—O5	−52.9	[−57.1, −46.6]	−53.7	[−56.3, −51.8]
C4—C5—O5—C1	55.9	[52.2, 58.6]	56.2	[54.5, 58.1]
C5—O5—C1—C2	−57.3	[−58.8, −55.6]	−56.0	[−57.5, −53.2]
O5—C1—C2—C3	55.7	[53.3, 59.5]	53.3	[46.6, 55.8]
O2—C2—C3—C4	−174.3	[−177.9, −172.2]	−173.3	[−176.0, −169.2]
O3—C3—C4—C5	171.7	[166.6, 176.6]	171.6	[170.2, 173.2]

TABLE XII

AM1 structural parameters of C_6O_5 subunits of the inclusion complex of β-CD and benzoic acid and those of β-CD; average values of each parameter and their ranges are given:



	β-CD + benzoic acid		β-CD	
	Ave.	Range	Ave.	Range
Bond length (Å)				
C1—C2	1.541	[1.538, 1.544]	1.541	[1.538, 1.544]
C2—C3	1.535	[1.534, 1.536]	1.535	[1.534, 1.536]
C3—C4	1.538	[1.536, 1.540]	1.538	[1.537, 1.540]
C4—C5	1.537	[1.535, 1.539]	1.537	[1.535, 1.538]
C5—C6	1.533	[1.532, 1.535]	1.533	[1.532, 1.534]
C5—O5	1.431	[1.430, 1.433]	1.431	[1.430, 1.432]
O5—C1	1.410	[1.405, 1.413]	1.411	[1.405, 1.414]
C1—O1	1.417	[1.414, 1.422]	1.417	[1.413, 1.421]
C2—O2	1.413	[1.407, 1.417]	1.413	[1.406, 1.416]
C3—O3	1.418	[1.416, 1.419]	1.417	[1.416, 1.418]
Bond angle (deg.)				
C1—C2—C3	110.2	[108.9, 112.1]	110.4	[109.6, 112.1]
C2—C3—C4	110.2	[108.6, 111.5]	110.0	[108.9, 111.2]
C3—C4—C5	110.9	[109.8, 112.0]	110.6	[110.1, 111.6]
C4—C5—C6	111.3	[110.0, 111.9]	111.3	[110.8, 111.8]
O5—C1—C2	111.9	[110.5, 113.4]	112.0	[111.0, 113.8]
O6—C6—C5	112.2	[111.6, 112.9]	112.0	[111.6, 112.5]
O2—C2—C3	111.2	[110.2, 112.6]	111.2	[110.3, 112.4]
O3—C3—C4	109.9	[106.8, 112.7]	109.9	[106.8, 112.0]
Dihedral angle (deg.)				
C1—C2—C3—C4	−52.5	[−56.1, −46.1]	−52.7	[−54.5, −46.7]
C2—C3—C4—C5	52.4	[50.3, 56.4]	53.3	[51.8, 55.6]
C3—C4—C5—O5	−52.6	[−56.0, −47.1]	−53.7	[−56.3, −51.8]
C4—C5—O5—C1	55.9	[52.6, 59.2]	56.2	[54.5, 58.1]
C5—O5—C1—C2	−56.7	[−58.9, −54.8]	−56.0	[−57.5, −53.2]
O5—C1—C2—C3	54.0	[47.2, 59.2]	53.3	[46.6, 55.8]
O2—C2—C3—C4	−173.0	[−175.8, −168.5]	−173.3	[−176.0, −169.2]
O3—C3—C4—C5	170.8	[167.8, 175.0]	171.6	[170.2, 173.2]

TABLE XIII

Heats of formation from 1SCF AM1 calculations for the β-CD with benzoic acid in the "tail-first" position with the guest moved along the *Z*-axis of host principal axis coordinate system stepwise.

Inclusion complexes	ΔH_f (kcal / mol)
β-CD + benzoic acid at +3 of the *Z*-axis	−1694.8
β-CD + benzoic acid at +2 of the *Z*-axis	−1706.3
β-CD + benzoic acid at +1 of the *Z*-axis	−1714.0
β-CD + benzoic acid at 0 of the *Z*-axis	−1716.9
β-CD + benzoic acid at −1 of the *Z*-axis	−1714.7
β-CD + benzoic acid at −2 of the *Z*-axis	−1713.8
β-CD + benzoic acid at −3 of the *Z*-axis	−1714.6

TABLE XIV

Heats of formation from 1SCF AM1 calculations for the β-CD with benzoic acid in the "head-first" position with the guest moved along the *Z*-axis of host principal axis coordinate system stepwise.

Inclusion complexes	ΔH_f (kcal / mol)
β-CD + benzoic acid at +3 of the *Z*-axis	−1724.3
β-CD + benzoic acid at +2 of the *Z*-axis	−1723.3
β-CD + benzoic acid at +1 of the *Z*-axis	−1724.2
β-CD + benzoic acid at 0 of the *Z*-axis	−1725.0
β-CD + benzoic acid at −1 of the *Z*-axis	−1723.1
β-CD + benzoic acid at −2 of the *Z*-axis	−1720.5
β-CD + benzoic acid at −3 of the *Z*-axis	−1718.0

that a greater portion of the benzene ring is in the host cavity, stabilization of the whole complex increases until it reaches the minimum-energy configuration. Then, when the carboxylic acid group is moved out from the primary alcohol, a greater portion of the benzene ring is out of the cavity, causing the increasing heats of formation. This may be due to the hydrophobic interaction between the host and guest.

In conclusion, we studied the inclusion complexes of α-CD and β-CD with benzoic acid and phenol in the "head-first" and "tail-first" positions. The driving forces for complex formation were investigated by examining combinations of different intermolecular interactions such as steric fit, dipole–dipole interaction, intramolecular hydrogen bonding, intermolecular hydrogen bonding, and the enthalpies of formation of host and guest molecules calculated at their geometries in the complex and at their optimized geometries.

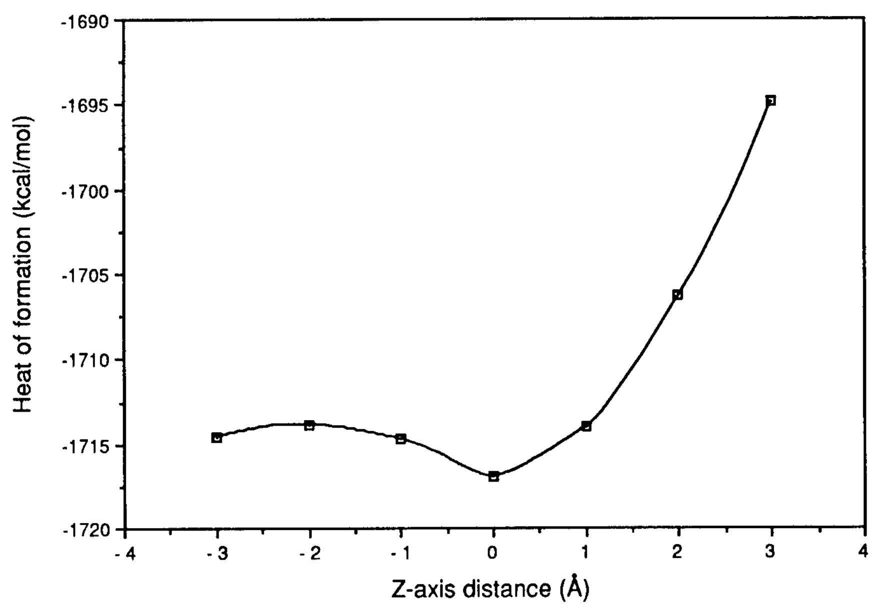


FIGURE 2. Plotting of 1SCF AM1 calculated ΔH_f vs. displacement of guest along *Z*-axis of host principal axis coordinate system for β-CD with benzoic acid in "tail first" position.

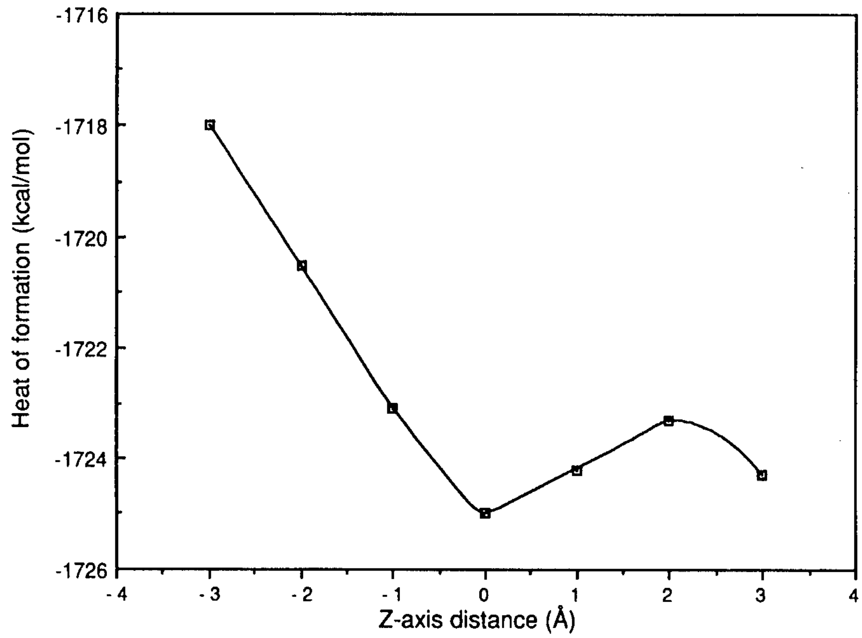


FIGURE 3. Plotting of 1SCF AM1 calculated ΔH_f vs. displacement of guest along Z-axis of host principal axis coordinate system for β-CD with benzoic acid in "head first" position.

References

1. A. Villiers, Compt. Rend. **112**, 536 (1891).
2. F. Schardinger, Wien. Klin. Wochschr. **17**, 207 (1904).
3. K. Freudenberg and R. Jacobi, Ann. Chem. **518**, 102 (1935).
4. W. Saenger, Angew. Chem., Int. Ed. Engl. **19**, 344 (1980).
5. (a) M. L. Bender and M. Komiyama, *Cyclodextrin Chemistry* (Springer: New York, 1978). (b) J. Szejtli, *Cyclodextrins and Their Inclusion Complexes* (Akademiai Kiado, Budapest, 1982). (c) S. Li and W. C. Purdy, Chem. Rev. **92**, 1457 (1992). (d) J. Szejtli, *Cyclodextrin Technology* (Kluwer, Dordrecht, 1988).
6. (a) K. Freudenberg and F. Cramer, Z. Naturforsch. Teil B **3**, 464 (1948). (b) K. Freudenberg and F. Cramer, Chem. Ber. **83**, 296 (1950).
7. (a) F. Cramer and H. Hettler, Naturwissenschaften **54**, 625 (1967). (b) W. Saenger, in *Inclusion Compounds*, J. L. Atwood, J. E. D. Davies, and D. D. MacNicol, Eds. (Academic Press, London, 1984), Vol. 2. (c) J. Szejtli, Kontakte (Darmstadt) (1), 31 (1988).
8. J. A. Hamilton and L. K. Steinrauf, Acta Crystallogr. Sect. B **24**, 1560 (1968).
9. D. J. Wood, F. E. Hruska, and W. Saenger, J. Am. Chem. Soc. **99**, 1735 (1977).
10. K. Flohr, R. M. Patton, and E. T. Kaiser, J. Am. Chem. Soc **97**, 1209 (1975).
11. (a) S. Dong and D. Zhang, Acta Chim. Sin. **46**, 335 (1988). (b) R. Ismin, C. Slam, and A. E. Kaifer, J. Org. Chem. **56**, 35 (1991).
12. R. L. Van Etten, J. F. Sebastain, G. A. Clowes, and M. L. Bender, J. Am. Chem. Soc. **89**, 3242 (1967).
13. F. Cramer, Angew. Chem. **68**, 115 (1956).
14. J. Nishijo and M. Nagai, J. Pharm. Sci. **80**, 58 (1991).
15. J. Cohen and J. L. Lach, J. Pharm. Sci. **52**, 132 (1963).
16. S. P. Jones, D. J. W. Grant, J. Hadgraft, and G. D. Parr, Acta Pharm. Technol. **30**, 213 (1984).
17. W.-Q. Tong, J. L. Lach, T.-F. Chin, and J. K. Guillory, Pharm. Res. **8**, 951 (1991).
18. I. Tabushi, Y. Kiyosuke, T. Sugiomoto, and K. Yamamura, J. Am. Chem. Soc. **100**, 916 (1978).
19. A. Orstan and J. B. A. Ross, J. Phys. Chem. **91**, 2735 (1987).
20. C. Jaime, J. Redondo, F. Sánchez-Ferrando, and A. Virgili, J. Mol. Struct. **248**, 317 (1991).
21. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, J. Am. Chem. Soc. **107**, 3902 (1985).
22. W.-Q. Tong, J. L. Lach, T.-F. Chin, and J. K. Guillory, Pharm. Res. **8**, 1307 (1991).
23. M. E. Amato, F. Djedaïni, G. C. Pappalardo, B. Perly, and G. Scarlata, J. Pharm. Sci. **81**, 1157 (1992).
24. A. S. Kostense, S. P. van Helden, and L. H. M. Janssen, J. Comput. Aid. Mol. Des. **5**, 525 (1991).
25. I. Tabushi, Y.-O. Kiyosuke, U. Sugimoto, and K. Yamamura, J. Am. Chem. Soc. **100**(3), 916 (1978).
26. Y. Matsui, Bull. Chem. Soc. Jpn. **55**(4), 1246 (1982).
27. D. W. Armstrong, T. J. Ward, R. D. Armstrong, and T. E. Beesley, Science **232**, 1132 (1986).
28. I. Tabushi and T. Mizutani, Tetrahedron **43**(7), 1439 (1987).

29. F. M. Menger and M. J. Sherrod, J. Am. Chem. Soc. **110**, 8606 (1988).
30. H.-J. Thiem, M. Brandl, and R. Breslow, J. Am. Chem. Soc. **110**, 8612 (1988).
31. C. A. Venanzi, P. M. Canzius, Z. Zhang, and J. D. Bunce, J. Comput. Chem. **10**(8), 1038 (1989).
32. M. J. Sherrod, Carbohydr. Res. **192**, 17 (1989).
33. A. F. D. deNamor, R. Traboulssi, and D. F. V. Lewis, J. Chem. Soc., Chem. Commun. 751 (1990).
34. M. Ohashi, K. Kasatani, H. Shinohara, and H. Sato, J. Am. Chem. Soc. **112**, 5824 (1990).
35. K. B. Lipkowitz, J. Org. Chem. **56**, 6357 (1991).
36. S. P. van Helden, B. P. van Eijck, and L. H. M. Janssen, J. Biomol. Struct. Dyn. **9**, 1269 (1992).
37. J. E. H. Koehler, W. Saenger, and W. F. van Gunsteren, J. Mol. Biol. **203**, 241 (1988).
38. J. E. H. Koehler, W. Saenger, and W. F. van Gunsteren, Eur. Biophys. J. **15**, 197 (1987).
39. J. E. H. Koehler, W. Saenger, and W. F. van Gunsteren, Eur. Biophys. J. **15**, 211 (1987).
40. M. Prabhakaran and S. C. Harvey, Biopolymers **26**, 1087 (1987).
41. J. E. H. Koehler, W. Saenger, and W. F. van Gunsteren, J. Biomol. Struct. Dyn. **6**(1), 181 (1988).
42. J. E. H. Koehler, W. Saenger, and W. F. van Gunsteren, Eur. Biophys. J. **16**, 153 (1988).
43. M. Kitagawa, H. Hoshi, M. Sakurai, Y. Inoue, and R. Chujo, Carbohydr. Res. **163**, C1 (1987).
44. M. Sakurai, M. Kitagawa, and H. Hoshi, Chem. Lett. 895 (1988).
45. M. Kitagawa, H. Hoshi, M. Sakurai, Y. Inoue, and R. Chujo, Bull. Chem. Soc. Jpn. **61**, 4225 (1988).
46. M. Sakurai, M. Kitagawa, H. Hoshi, Y. Inoue, and R. Chujo, Bull. Chem. Soc. Jpn. **62**, 2067 (1989).
47. N. S. Bodor, M.-J. Huang, and J. D. Watts, J. Pharm. Sci. **84**, 330 (1995).
48. N. S. Bodor and M.-J. Huang, J. Pharm. Sci. **81**, 272 (1992).
49. J. B. Pedley and G. Rylance, *Sussex-N. P. L. Computer Analysed Thermochemical Data: Organic and Organometallic Compounds* (Sussex University, 1977).
50. C. Van Hooidonk and J. C. A. E. Breebaart-Hansen, Rec. Trav. Chim. **91**, 958 (1972).
51. E. Siimer, M. Kurvits, and A. Kostner, Thermochim. Acta **116**, 249 (1987).
52. E. Siimer and M. Kurvits, Thermochim. Acta **140**, 161 (1989).
53. M. Kamiya, S. Mitsuhashi, M. Makino, and H. Yoshioka, J. Phys. Chem. **96**, 95 (1992).
54. G. Marconi, S. Monti, B. Mayer, and G. Köhler, J. Phys. Chem. **99**, 3943 (1995).
55. G. A. Jeffrey and W. Saenger, *Hydrogen Bonding in Biological Structures* (Springer-Verlag, Berlin, 1991).
56. Th. Steiner and W. Saenger, Acta Crystallogr., Sect. B **8**, 819 (1992).
57. Th. Steiner and W. Saenger, J. Am. Chem. Soc. **115**, 4540 (1993).
58. Th. Steiner and W. Saenger, Carbohydr. Res. **259**, 1 (1994).
59. W. Saenger, in *Inclusion Compounds*, Vol. 2, *Structural Aspects of Inclusion Compounds Formed by Organic Host Lattices*, J. L. Atwood, J. E. D. Davies, and D. D. MacNicol (Academic Press, London, 1984).

List of Participants

WILLIAM H. ADAMS
Rutgers University
Department of Chemistry
P.O. Box 939
Piscataway, NJ 08855-0939
USA
Phone: 908-445-3758
Fax: 908-445-5312
EMail: adams@rutchem.rutgers.edu

ESTHER AGACINO
Universidad Nacional Autónoma de Mexico
Centro de Investigaciones Teoricas
Avenida 10 de Mayo s/n
Cuautitlan Izcalli, Mexico 54740
Mexico
Phone: 5-623-20-37
Fax: 5-623-20-33
EMail: eav@hp.fciencias.unam.mx

JANOS G. ANGYAN
Universite Henri Poincare
Laboratoire de Chimie Theorique
B. P. 239
Vandoeuvre, Les Nancy Cedex F-54506
France
Phone: 33-3-83-91-25-29
Fax: 33-3-83-91-25-30
EMail: angyan@lctn.u-nancy.fr

JOHAN ÅQVIST
Uppsala University
Department of Molecular Biology
P.O. Box 590
Uppsala S-751 24
Sweden
Phone: 46-18-174109
Fax: 46-18-536971
EMail: aqvist@xray.bmc.uu.se

PETER B. ARMENTROUT
University of Utah
Department of Chemistry
Salt Lake City, UT 84112
USA
Phone: 801-581-7885
Fax: 801-581-8433
EMail: armentrout@chemistry.utah.edu

GUSTAVO A. ARTECA
Laurentian University
Chemistry and Biochemistry
Ramsey Lake Road
Sudbury, Ontario P3E 2C6
Canada
Phone: 705-675-1151
Fax: 705-675-4844
EMail: gustavo@nickel.laurentian.ca

OSMAN ATABEK
CNRS - Universite Paris - Sud
Lab Photophysique Moleculaire
BAT. 213 Campus D'Orsay
Orsay 91405
France
Phone: 33-1-69416132
Fax: 33-1-69156777
EMail: osman@atabek.ppm.u-psud.fr

KYOUNG K. BAECK
University of Florida
Quantum Theory Project
362 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: baeck@knusun.kangnung.ac.kr

International Journal of Quantum Chemistry, Vol. 65: 1153–1177 (1997)
© 1997 John Wiley & Sons, Inc.
CCC 0020-7608 / 97 / 061153-25

DIRK BAKOWIES
University of California at San Francisco
Dept. of Pharm. Chem.
Box 0446
San Francisco, CA 94143-0446
USA
Phone: 415-476-4026
Fax: 415-476-0688
EMail: bakowies@clg.ucsf.edu

ANNA BALKOVA
University of Florida
Quantum Theory Project
P.O. Box 118435
Gainesville, FL 32611
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: balkova@qtp.ufl.edu

ANDRE D. BANDRAUK
University of Sherbrooke
Chimie-Sciences
2500 Blvd. Universite
Sherbrooke, Quebec J1K 2R1
Canada
Phone: 819-821-7098
Fax: 819-821-8017
EMail: bandrauk@planck.chimie.usherb.ca

GABRIELA BARREIRO
Universidade Federal do Rio de Janeiro
Inst. de Quimica
Rua Maestro Francisco Braga 585, Apto. 202
Rio de Janeiro 22041-070
Brazil

ANDRES BARRIOS
Florida International University
Dept. of Electrical & Computer Eng.
Miami, FL 33199
USA
Phone: 305-348-3017
Fax: 305-348-3707
EMail: hagmann@eng.fiu.edu

RODNEY J. BARTLETT
University of Florida
Quantum Theory Project
381 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: bartlett@qtp.ufl.edu

DONALD R. BECK
Michigan Technological University
Dept. of Physics
Houghton, MI 49931
USA
Phone: 906-487-2019
Fax: 906-487-2933
EMail: donald@phy.mtu.edu

SULLIVAN BECK
University of Florida
Quantum Theory Project
345 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: beck@qtp.ufl.edu

THOMAS BECK
University of Cincinnati
Dept. of Chemistry
ML172
Cincinnati, OH 45221
USA
Phone: 513-556-4886
Fax: 513-556-9239
EMail: tlb@moe.che.uc.edu

CHRIS BENDER
Albert Einstein College of Medicine
Molecular Pharmacology
1300 Morris Park Ave.
Bronx, NY 10461
USA
Phone: 718-430-2175
Fax: 718-430-8922
EMail: bender@spin.aecom.yu.edu

ATILLA BÉRCES
Steacie Institute for Molecular Science
100 Susses Drive, Room 1150
Ottawa, Ontario K1A 0R6
Canada
Phone: 613-990-0964
Fax: 613-954-5242
EMail: attila@ned1.sims.nrc.ca

JERZY BERNHOLC
North Carolina State University
Dept. of Physics
Box 8202
Raleigh, NC 27695-8202
USA
Phone: 919-515-3126
Fax: 919-515-7331
EMail: bernholc@ncsu.edu

R. STEPHEN BERRY
University of Chicago
Dept. of Chemistry
5735 South Ellis Avenue
Chicago, IL 60637-1403
USA
Phone: 312-702-7021
Fax: 312-702-0805
EMail: berry@rainbow.uchicago.edu

VIRINEYA BERTIN
Universidad Autónoma Metropolitana - Iztapalapa
Dept. de Quimica CIB
AP 55-534
Mexico, DF CP 09340
Mexico
Phone: 52-5-724-4668
Fax: 52-5-724-4666
EMail: neya@xanum.uam.mx

DAVID BEVERIDGE
Wesleyan University
Department of Chemistry
Hall-Atwater Laboratories
Middletown, CT 06457
USA
Phone: 860-685-2575
Fax: 860-685-2211
EMail: dbeveridge@wesleyan.edu

APURBA BHATTACHARJEE
Walter Reed Army Institute of Research
Dept. of Pharmacology
Washington, DC 20307-5100
USA
Phone: 301-295-7085
Fax: 301-295-7755
EMail: apurba_bhattacharjee@wrsmtp-ccmail.army.mil

RICARDO BICCA DE ALENCASTRO
Universidade Federal de Rio de Janeiro
Inst. de Quimica
Bloco A-CT Sala 622, Cidade Univ.
Rio de Janeiro, RJ 21949-900
Brazil
Phone: 55-021-590-3544
Fax: 55-021-290-4746
EMail: bicca@iq.ufrj.br

DAVID M. BISHOP
University of Ottawa
Dept. of Chemistry
P.O. Box 450, STN A
Ottawa, Ontario K1N 6N5
Canada
Phone: 613-562-5181
Fax: 613-562-5170
EMail: dbishop@oreo.uottawa.ca

FRANK E. BLANEY
Computational Chemistry
Smithkline Beecham
New Frontiers Science Park (N) Third Avenue
Harlow, Essex CM19 5AW
UK
Phone: 44-1279-622143
Fax: 44-1279-622348
EMail: frank_e_blaney@sbphrd.com

JONATHAN C. BOETTGER
Los Alamos National Lab.
Group T-1 MS B221
Los Alamos, NM 87545
USA
Phone: 505-667-7483
Fax: 505-665-5757
EMail: jn@lanl.gov

ALEXANDER L. BOLDYREV
University of Utah
Dept. of Chemistry
Salt Lake City, UT 84112-1194
USA
Phone: 801-581-7445
Fax: 801-581-8433
EMail: boldyrev@chemistry.utah.edu

EDWARD A. BOUDREAUX
University of New Orleans
Dept. of Chemistry
Lake Front
New Orleans, LA 70148
USA
Phone: 504-286-6311
Fax: 504-286-6860

AHMED BOUFERGUENE
Florida A & M University
Dept. of Physics
205 Jones Hall
Tallahassee, FL 32307
USA
Phone: 904-599-3470
Fax: 904-599-3577
EMail: boufer@cennas.nhmfl.gov

JOEL M. BOWMAN
Emory University
Dept. of Chemistry
1515 Pierce Drive
Atlanta, GA 30322
USA
Phone: 404-727-6592
Fax: 404-727-6628
EMail: bowman@euch3g.emory.edu

ERKKI BRÄNDAS
Uppsala University
Dept. of Quantum Chemistry
Box 518
Uppsala S-75120
Sweden
Phone: 46-18-183263
Fax: 46-18-502402
EMail: erkki@kvac.uu.se

RIA BROER
University of Groningen
Dept. of Chemical Physics
Nijenborgh 4
Groningen, AG 9747
The Netherlands
Phone: 31-50-634374

ANDERS BROO
University of Florida
Quantum Theory Project
P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: broo@qtp.ufl.edu

LUIS BRUNO-BLANCH
Universidad Nacional de La Plata
Depto Ciencias Biological
CC 243
La Plata, Buenos Aires 1900
Argentina
Phone: 54-21-210784 EXT 41
Fax: 54-21-254533
EMail: lbb@nahuel.biol.unlp.edu.ar

JOHNNY BUSTAD
University College of Gavle
Science
S-801 76
Gavle, Sweden
Sweden
Phone: 46-26-648758
EMail: jbd@hgs.se

SYLVIO CANUTO
Universidad de São Paulo
Instituto de Fisica
CPX 66318
São Paulo 05315-970
Brazil
Phone: 55-11-818-6983
Fax: 55-11-818-6831
EMail: canuto@if.usp.br

CRISTIÁN CÁRDENAS-LAILHACAR
University of Florida
Quantum Theory Project
P.O. Box 118435, 345 Williamson Hall
Gainesville, FL 32611-8435
USA
Phone: 352-392-6365
Fax: 352-392-8722
EMail: cardenas@qtp.ufl.edu

RAQUEL CASTILLO
Universitat Jaune I
Dept. de Ciencias Experimentales
Apartat 242
Castellon 12080
Spain
Phone: 34-64-345700
Fax: 34-64-345654
EMail: rcastill@vents.uji.es

SIDONIO CASTILLO
Universidad Autonoma Metropolitana-
Azcapotzalco
Ciencias Basicas
Av. San Pablo #180
Mexico, DF 02200
Mexico
Phone: 52-5-724-4218
Fax: 52-5-723-5940
EMail: sca@hp9000a1.uam.mx

EDUARDO A. CASTRO
Universidad Nacional de La Plata
Dept. of Chemistry
Calle 47 Y 115 C C 962
La Plata, Buenos Aires 1900
Argentina
Phone: 54-21-214037
Fax: 54-21-259485
EMail: castro@biol.unlp.edu.ar

MIGUEL CASTRO
Universidad Nacional Autónoma de Mexico
Dept. de Fisica y Quimica Teorica
Mexico, DF 04510
Mexico
Phone: 52-5-622-3783
Fax: 52-5-616-2010
EMail: castro@papalotl.pquim.unam.mx

CARY F. CHABALOWSKI
U.S. Army Research Lab
AMSRL-WT-PC
Aberdeen Prov. Gr., MD 21005-5066
USA
Phone: 410-278-6095
Fax: 410-278-6150
EMail: cary@arl.mil

BENOIT CHAMPAGNE
Universitaires Notre-Dame de la Paix, à Namur
CTA Lab
Rue de Bruxelles 61
Namur B-5000
Belgium
Phone: 32-81-724557
Fax: 32-81-724530
EMail: benoit.champagne@fundp.ac.be

YAT-PING CHANG
National Taiwan University
Dept. of Chemistry
Room 203
Taipei
Taiwan, ROC
Phone: 886-2-3630231-2326
Fax: 886-2-3636359
EMail: ypchang@rs350.ch.ntu.edu.tw

JOSEPH R. CHAVEZ
Air Force
6221 Cole Lane
Albuquerque, NM 87105
USA
Phone: 505-846-6456
Fax: 505-846-2290
EMail: chavezj@plk.af.mil

JIAN CHEN
University of Florida
Dept. of Physics
215 Williamson Hall
Gainesville, FL 32611
USA
Phone: 352-392-0521
EMail: jchem@physics.ufl.edu

HAI-PING CHENG
University of Florida
Quantum Theory Project
371 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: cheng@qtp.ufl.edu

DANIEL M. CHIPMAN
University of Notre Dame
Radiation Laboratory
Notre Dame, IN 46556-0579
USA
Phone: 219-631-5562

SAN-YAN CHU
National Tsing Hua University, Taiwan
Department of Chemistry
Hsinchu
Taiwan
Phone: 866-35-721634
Fax: 886-35-711082
EMail: sychu@chem.nthu.edu.tw

PAUL CHUN
University of Florida
Dept. of Biochem. & Mol. Biol.
Box 100245
Gainesville, FL 32610-0245
USA
Phone: 352-392-3356
Fax: 352-392-2953
EMail: pwchun@pine.circa.ufl.edu

JERZY CIOSLOWSKI
Florida State University
Dept. of Chemistry
Tallahassee, FL 32306-3006
USA
Phone: 904-644-8274
Fax: 904-644-8281
EMail: jerzy@kyoko.chem.fsu.edu

JAMES COFFIN
IBM
1505 LBJ Freeway, 5th Floor
Computational Chemistry
Dallas, TX 75234
USA
Phone: 214-280-3960
Fax: 214-280-2779

ROBERT N. COMPTON
University of Tennessee
Dept. of Chemistry and Physics
401 Buehler Hall
Knoxville, TN 37996
USA
Phone: 423-574-6233
Fax: 423-974-3454
EMail: ahd@ornc.gov

PERE CONSTANS-NIEREA
University of Girona
Inst. of Computational Chemistry
Albereda 3-5
Girona 17921
Spain
Phone: 34-72-418357
Fax: 34-72-418357
EMail: pere@infern.udg.es

DAVID COOPER
University of Liverpool
Dept. of Chemistry
P.O. Box 147
Liverpool L69 3BX
UK
Phone: 44-151-794-3532
Fax: 44-151-794-3588
EMail: dlc@liv.ac.uk

GRANT COOPER
International Physics Health & Energy, Inc.
2260 W. Holcombe Blvd.
Houston, TX 77030
USA
Phone: 713-523-5459
Fax: 713-522-2607

JOAO MANUEL CORDEIRO
Universidad Estadual Paulista
Dept. Fisica e Quimica
Av. Brasil 56, P.O. Box 31
Ilha Solteira, SP 15378-000
Brazil
Phone: 55-18-762-3850
Fax: 55-18-762-4868
EMail: cordeiro@isl000.uesp.ansp.br

MARSHALL CORY
University of Florida
Quantum Theory Project
380 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: cory@qtp.ufl.edu

KALINE COUTINHO
Universidad de São Paulo
Instituto de Fisica
CP 66318
São Paulo 05315-970
Brazil
Phone: 55-11-818-6983
Fax: 55-11-818-6831
EMail: kaline@if.usp.br

MAURICIO COUTINHO-NETO
University of Florida
Quantum Theory Project
343 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: coutinho@qtp.ufl.edu

T. DANIEL CRAWFORD
University of Georgia
Dept. of Chemistry
1001 Cedar Street
Athens, GA 30602
USA
Phone: 706-542-7738
Fax: 706-542-0406
EMail: crawdad@otanes.ccqc.uga.edu

ARMANDO CRUZ
Instituto Mexicano del Petroleo
AP 14-805
Mexico, DF 07730
Mexico
Phone: 52-5-368-5911 EXT 20270
Fax: 52-5-567-2927
EMail: armando@briseida.sgi.imp.mx

THOMAS CUNDARI
Memphis State University
Dept. of Chemistry
3744 Walker
Memphis, TN 38152
USA
Phone: 901-678-2629
Fax: 901-678-3447
EMail: tcundari@cc.memphis.edu

HERBERT DA COSTA
University of Florida
Quantum Theory Project
P.O. Box 118435, 348 Williamson Hall
Gainesville, FL 32611-8435
USA
Phone: 352-392-7184
Fax: 352-392-8722
EMail: dacosta@qtp.ufl.edu

JOAQUIM DA MOTTA NETO
University of Florida
Quantum Theory Project
348 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: quim@qtp.ufl.edu

GURU P. DAS
Wright Laboratory
MLBP
WPAFB, OH 45433
USA
Phone: 937-429-2307
Fax: 937-255-9147
EMail: dasgp@picard.ml.wpafb.af.mil

ROBIN H. DAVIES
University of Wales at Cardiff
Welsh School of Pharmacy
King Edward VII Ave.
Cardiff, Wales CF1 3XF
UK
Phone: 44-1222-874000 EXT 5830
Fax: 44-1222-874149

WAGNER B. DE ALMEIDA
University of Florida
Quantum Theory Project
P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: wagner@qtp.ufl.edu

DAN DE KEE
University of Florida
Quantum Theory Project
P.O. Box 118435
Gainesville, FL 32611
USA
Phone: 352-392-9306
Fax: 352-392-8722
EMail: dekee@qtp.ufl.edu

GLENISSON DE OLIVEIRA
Indiana University - Purdue University Indianapolis
Dept. of Chemistry
402 North Blackford St.
Indianapolis, IN 46202
USA
Phone: 317-274-5342
Fax: 317-274-4701
EMail: glen@chem.iupui.edu

DOMINIQUE DEHARENG
Universite de Liege
Institut de Chimie
B6 Sart Tilman
Liege B-4000
Belgium
Phone: 32-4-3663-499
Fax: 32-4-3663-364
EMail: domi@iris.cip.ulg.ac.be

JANET DEL BENE
Youngstown State University
Dept. of Chemistry
One University Plaza
Youngstown, OH 44555
USA
Phone: 330-742-3466
Fax: 330-742-1579
EMail: fro42008@ysub.ysu.edu

JUAN-CARLOS DEL VALLE
Florida State University
Tallahassee, FL
USA
Phone: 904-644-6452

ERIK DEUMENS
University of Florida
Quantum Theory Project
359 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: deumens@qtp.ufl.edu

GEORGES DIVE
University of Liege
Institut de Chimie
B6 Sart Tilman
Liege B-4000
Belgium
Phone: 32-4-366-3499
Fax: 32-4-366-3364
EMail: gd@iris.cir.ul6.ac.be

RUSSELL DRAGO
University of Florida
CLB 408
Gainesville, FL 32611
USA
Phone: 352-392-6043

JAMES DUFTY
University of Florida
Dept. of Physics
215 Williamson Hall, P.O. Box 118440
Gainesville, FL 32611
USA
Phone: 352-392-6693
Fax: 352-392-0524
EMail: dufty@phys.ufl.edu

MICHEL DUPUIS
Pacific Northwest National Laboratory
Mail Stop K1-90
P.O. Box 999, 906 Battelle Boulevard
Richland, WA 99352
USA
Phone: 509-375-6784
Fax: 509-375-6631
EMail: m_dupuis@pnl.gov

MOSTAFA A. EL-SAYED
Georgia Institute of Technology
School of Chemistry
Boggs Building
Atlanta, GA 30332-0400
USA
Phone: 404-894-0294
Fax: 404-894-0294
EMail: mostafa.el-sayed@chemistry.gatech.edu

GÉRHARD G. EMCH
University of Florida
Mathematics Dept.
P.O. Box 118105
Gainesville, FL 32611-8105
USA
Phone: 352-392-0281 EXT 274
Fax: 352-392-8357
EMail: gge@math.ufl.edu

GUILLERMINA L. ESTIU
Universidad Nacional de La Plata
Dept. Quimica
Calle 47 Y 115 CC962
La Plata, Buenos Aires 1900
Argentina
Phone: 54-21-210784 EXT 41
Fax: 54-21-259485
EMail: estiu@nahuel.biol.unlp.edu.ar

TOM EVANS
University of Utah
Dept. of Chemistry
Salt Lake City, UT 84112
USA
Phone: 801-581-5465
Fax: 801-581-8433
EMail: evans@mercury.chem.utah.edu

GREGORY EZRA
Cornell University
Dept. of Chemistry
Baker Laboratory
Ithaca, NY 14853
USA
Phone: 607-255-3949
Fax: 607-255-4137
EMail: gse1@cornell.edu

WILLIAM FARIS
National Science Foundation
Math Science
4201 Wilson Blvd.
Arlington, VA 22230
USA
Phone: 703-306-1997

ANTONIO FERREIRA
University of Memphis
Department of Chemistry
Memphis, TN 38152
USA
Phone: 901-678-4429
Fax: 901-678-3447
EMail: amferreira@cc.memphis.edu

KIM FERRIS
Pacific Northwest Labs.
Materials & Chemical Sci. Ctr.
P.O. Box 999 MS K2-44
Richland, WA 99352
USA
Phone: 509-375-3754

SIGHART FISCHER
Technische Universität München
Theoretische Physik T 38
Boltzmann Str.
Garching B München 85747
Germany
Phone: 49-89-289-12373
Fax: 49-89-289-12444
EMail: jill@jupiter.t30.physik.tu-muenchen.de

EUGENE FLEISCHMANN
Q-Chem Inc.
317 Whipple Street
Pittsburgh, PA 15218
USA
Phone: 609-896-3942
Fax: 609-896-1244
EMail: gene@q-chem.com

GERNOT FRENKING
Phillips Universität Marburg
Dept. of Chemistry
Hans-Meerwein-Strasse
Marburg, Hessen 35032
Germany
Phone: 49-6421-285563
Fax: 49-6421-282189
EMail: frenking@ps1515.chemie.uni-marburg.de

PIOTR FROELICH
Uppsala University
Dept. of Quantum Chemistry
Box 518
Uppsala 751 20
Sweden
Phone: 46-18-183262
Fax: 46-18-502402
EMail: piotr@kvac.uu.se

J. WILLIAM GADZUK
National Institute of Standards and Technology
Gaithersburg, MD 20899
USA
Phone: 301-975-2548
Fax: 301-926-6689
EMail: gadzuk@nist.gov

ANTHONY F. GARITO
University of Pennsylvania
Dept. of Physics and Astronomy
209 South 33 Street
Philadelphia, PA 19104
USA
Phone: 215-898-5810
Fax: 215-898-2010
EMail: garito@afgarito.physics.upenn.edu

JOSE M. GOMEZ LLORENTE
University of La Laguna
Fisica Fundamental
La Laguna 38205
Spain
Phone: 34-22-635448
Fax: 34-22-256973
EMail: jmgomez@ull.es

LIONEL GOODMAN
Rutgers University
Dept. of Chemistry
New Brunswick, NJ 08903
USA
Phone: 908-445-2603
Fax: 908-445-5312
EMail: goodman@rutchem.rutgers.edu

LEONID GORB
National Academy of Sciences of Ukraine
Inst. of Water and Colloid Chemistry
2, Skovorada Str.
Kiev 252680
Ukraine
Phone: 380-44-444-0196
Fax: 380-44-452-0276
EMail: glg@icck.freenet.kiev.ua

ALEX GREEN
University of Florida
ICAAS
P.O. Box 112050
Gainesville, FL 32611-2050
USA
Phone: 352-392-2001
Fax: 352-392-2027

CRISTIANO R. W. GUIMARÃES
Universidade Federal de Rio de Janeiro
Inst. de Quimica
Rua Prof Valadares 193, Apto. 901 - Grajau
Rio de Janeiro, RJ 20561-020
Brazil
Phone: 55-21-268-4837
Fax: 55-21-290-4746
EMail: gis@aix143.iq.ufrj.br

GENNADY L. GUTSEV
University of Florida
Quantum Theory Project
370 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: gutsev@qtp.ufl.edu

STEVEN R. GWALTNEY
University of Florida
Quantum Theory Project
345 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-6365
Fax: 352-392-8722
EMail: gwaltney@qtp.ufl.edu

MASAHIKO HADA
Kyoto University
Dept. of Synthetic Chemistry
Kyoto Sakyo-Ku, Yoshida-Hon-Machi 606
Japan
Phone: 81-75-5663
Fax: 81-75-5910
EMail: hada@synchem.kyoto-u.ac.jp

MARK J. HAGMANN
Florida International University
Dept. of Elect. & Computer Eng.
Miami, FL 33199
USA
Phone: 305-348-3017
Fax: 305-348-3707
EMail: hagmann@eng.fiu.edu

AAGE HANSEN
University of Copenhagen
Dept. of Chemistry
Universitetsparken 5
Copenhagen DK-2100
Denmark
Phone: 45-353-20284
Fax: 45-353-20299
EMail: aage@kl4ibim.ki.ku.dk

FRANK HARRIS
University of Utah
Dept. of Chemistry
Salt Lake City, UT 84112
USA
Phone: 801-581-8445
Fax: 801-585-3207
EMail: harris@dirac.chem.utah.edu

SERGIO ALEJANDRO HASSAN
Universidad Estadual de Campinas (UNICAMP)
Instituto de Fisica Gleb Wataghin
13083-970
Campinas, São Paulo
Brazil
Phone: 55-192-397254 or 392424
Fax: 55-192-393-127
EMail: mago@ifi.unicamp.br

JOHN HEAD
University of Hawaii
Dept. of Chemistry
2545 The Mall
Honolulu, HI 96822
USA
Phone: 808-956-5787
Fax: 808-956-5908
EMail: john@gold.chem.hawaii.edu

ARTHUR HEBARD
University of Florida
Dept. of Physics
215 Williamson Hall, P.O. Box 118440
Gainesville, FL 32611-8440
USA
Phone: 352-392-8842
Fax: 352-392-8586
EMail: afh@phys.ufl.edu

MAGNUS HEDSTRÖM
University of Florida
Quantum Theory Project
362 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-6973
Fax: 352-392-8722
EMail: hedstrom@qtp.ufl.edu

NIELS ENGHOLM HENRIKSEN
Technical University of Denmark
Dept. of Chemistry
Building 207
Lingby DK-2800
Denmark
Phone: 45-45252029
Fax: 45-45881639
EMail: neh@hcocvx.uni-c.dk

YA-WEN HSIAO
University of Florida
Quantum Theory Project
348 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: hsiao@qtp.ufl.edu

ZHENMING HU
Kyoto University
Dept. of Synthetic Chemistry and Biological Chemistry
Sakyo-Ku, Kyoto 606-01
Japan
Phone: 81-75-753-5659
Fax: 81-75-753-5910
EMail: hu@sbchem.kyoto-u.ac.jp

MING-JU HUANG
University of Florida
Center for Drug Discovery
Box 100497
Gainesville, FL 32610
USA
Phone: 904-392-8186
Fax: 904-392-8589
EMail: mjhuang@pine.circa.ufl.edu

YASUYUKI ISHIKAWA
University of Puerto Rico
Chemistry Department
P.O. Box 23346
San Juan, Puerto Rico 00931-3346
USA
Phone: 809-764-0000 EXT 7399
Fax: 809-751-0625
EMail: ishikawa@upracd.upr.clu.edu

ALBERT ISRAEL
Northwestern University
Chemistry
2145 Sheridan Road
Evanston, IL 60208
USA
Phone: 847-491-3423
EMail: albert@chem.nwu.edu

SAUL JACCHIERI
Instituto Adolfo Lutz
Divisao de Biologia
Av. Dr. Arnaldo 351 10 Andar
São Paulo, SP 01246902
Brazil
Phone: 55-11-30610111 EXT 215
Fax: 55-11-2750277

DENIS JACQUEMIN
University of Florida
Quantum Theory Project
P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-395-1597
Fax: 352-392-8722
EMail: jacquemin@qtp.ufl.edu

MORTEN O. JENSEN
Copenhagen University
H. C. Orsted Institute
Universitetsparken 5
Kobenhavn O DK-2100
Denmark
Phone: 45-353-20286
EMail: morten@kl4aix.ki.ku.dk

WILLIAM JORGENSEN
Yale University
Chemistry Department
225 Prospect Street
New Haven, CT 06511-8118
USA
Phone: 203-432-3916

ALICIA H. JUBERT
Universidad Nacional de La Plata
Facultad de Ciencias Exactas
CC 962
La Plata 1900
Argentina
Phone: 54-21-259485
Fax: 54-21-259485
EMail: jubert@nahuel.biol.unlp.edu.ar

PETER JURS
Penn State University
Dept. of Chemistry
152 Davey Lab.
University Park, PA 16802
USA
Phone: 814-865-3739
Fax: 814-865-3314
EMail: pcj@psu.edu

SABRE KAIS
Purdue University
Chemistry Department
1393 Brown Building
West Lafayette, IN 47907
USA
Phone: 317-494-5965
Fax: 317-494-0239
EMail: sabre@salam.chem.purdue.edu

JEAN KARLE
Walter Reed Army Institute of Research
Pharmacology
Division of Experimental Therapeutics
Washington, DC 20307-5100
USA
Phone: 301-295-7191
Fax: 301-295-7755
EMail: dr.-jean-karle@wrsmtp-ccmail.army.mil

LIST OF PARTICIPANTS

JEROME KARLE
Naval Research Laboratory
Laboratory for the Structure of Matter
Code 6030 Naval Research Laboratory
Washington, DC 20375-5341
USA
Phone: 202-767-2665
Fax: 202-767-0953
EMail: williams@harker.nrl.navy.mil

SASHI KARNA
U.S. Air Force Phillips Laboratory, VTM
Space Mission Technologies Div.
3550 Aberdeen Ave. SE
Kirtland AFB, NM 87117-5776
USA
Phone: 505-853-3158
Fax: 505-846-2290
EMail: karnas@plk.af.mil

MICHAEL KASHA
Florida State University
Inst. of Molec. Biophys.
452 Molecular Biophysics
Tallahassee, FL 32306
USA
Phone: 904-644-6452

HIROYUKI KAWABE
Kinjo College
Secretarial Department 200
Matto, Ishikawa 924
Japan
Phone: 81-762-76-4411
Fax: 81-762-75-4183
EMail: kawabe@kinjo.ac.jp

MICHAEL E. KELLMAN
University of Oregon
Dept. of Chemistry
Eugene, OR 97403
USA
Phone: 541-346-4196
Fax: 541-346-4643
EMail: kellman@oregon.uoregon.edu

CHARLES BRIAN KELLOGG
University of Georgia
Chemistry
4th Old Chemistry
Athens, GA 30602
USA
Phone: 404-542-2067
Fax: 404-542-0406

ANNE-MARIE KELTERER
Technical University of Graz
Physical and Theoretical Chemistry
Rechbauer Street 12
Graz 8010
Austria
Phone: 43-316-873-8727
EMail: kelterer@ptc.tu-graz.ac.at

JAMES W. KING
Foundation for Chemistry
P.O. Box 116
Balsam, NC 28707-0116
USA
Phone: 704-452-7570
Fax: 704-452-5432
EMail: jwking@sprynet.com

BERNARD KIRTMAN
University of California, Santa Barbara
Dept. of Chemistry
Santa Barbara, CA 93106
USA
Phone: 805-893-2217
Fax: 805-893-4120
EMail: kirtman@sbmm1.ucsb.edu

DOUGLAS J. KLEIN
Texas A & M University
Marine Science Department
P.O. Box 1675
Galveston, TX 77553-1675
USA
Phone: 409-740-4512
Fax: 409-740-4429

MARTIN KLESSINGER
Universität Münster
Organisch-Chemisches Institut
Correnstr. 40
Muenster D-48149
Germany
Phone: 49-251-8333-241
Fax: 49-251-8339-772
EMail: klessim@uni-muenster.de

GILLES KLOPMAN
University of Florida
Chemistry
P.O. Box 118435
Gainesville, FL 32611
USA
Phone: 352-392-6711

TAKAO KOBAYASHI
Osaka University
Grad. School of Science Chemistry
Toyonaka, Osaka 560
Japan
Phone: 81-6-850-5405
Fax: 81-6-843-7744
EMail: tkoba@chem.sci.osaka-u.ac.jp

PETER KOLLMAN
University of California
Pharmaceutical Chemistry Dept.
School of Pharmacy
San Francisco, CA 94143-0446
USA
Phone: 415-476-4637
Fax: 415-476-0688
EMail: pak@cgl.ucsf.edu

ANATOLI KORKIN
University of Florida
Quantum Theory Project
362 WM Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-8113
Fax: 352-392-8722
EMail: korkin@qtp.ufl.edu

SVETLANA KOTOCHIGOVA
National Institute of Standards and Technology
Dept. of Physics
A253 Physics Building, Quince Orchard Road
Gaithersburg, MD 20899
USA
Phone: 301-975-3732
Fax: 301-975-3038
EMail: svetlana@bruce.nist.gov

ATILLA KOVÁCS
Budapest Technical University
Inst. for General & Analytical Chemistry
Srt. Gellert ter 4
Budapest H-1521
Hungary
Phone: 36-1-463-3414
Fax: 36-1-463-4052
EMail: kovacs@ch.bme.hu

JEFFREY L. KRAUSE
University of Florida
Quantum Theory Project
372 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: krause@qtp.ufl.edu

CARLOS KUBLI-GARFIAS
National Autonomous University of Mexico
Lab. of Hormonal Chemistry
Apdo Postal 70-469
Mexico, DF 04511
Mexico
Phone: 525-6-223815
Fax: 525-5-500048
EMail: kubli@servidor.unam.mx

HENRY KURTZ
University of Memphis
Dept. of Chemistry
Memphis, TN 38152-6060
USA
Phone: 901-678-4414
Fax: 901-678-3447
EMail: kurtz@cc.memphis.edu

JANOS LADIK
Numberg Universität-Erlangen
Inst. of Theoretical Chemistry
Egerlandstrasse 3
Erlangen D-97058
Germany
Phone: 49-9131-857766
Fax: 49-9131-857736
EMail: ladik@pctc.chemie.uni-erlangen.de

WILLIAM LAIDIG
Procter & Gamble Co.
Miami Valley Laboratories, P.O. Box 538707
Cincinnati, OH 45253-8707
USA
Phone: 513-627-2857

DAVID LANGRETH
Rutgers University
Dept. of Physics & Astronomy
P.O. Box 849
Piscataway, NJ 08855-0849
USA
Phone: 908-445-3882
Fax: 908-445-4400
EMail: langreth@physics.rutgers.edu

PIERRE LEBRETON
University of Illinois at Chicago
Dept. of Chemistry
845 W. Taylor St.
Chicago, IL 60607-7061
USA
Phone: 312-996-5431
Fax: 312-996-0431
EMail: lebreton@uic.edu

ROLAND LEFEBVRE
Univ. de Paris SUD
Laborataire de Photophysique Moleculaire
Lab. Photphys. Molbatiment 213
Orsay F-09405
France
Phone: 33-1-63416533
Fax: 33-1-69416777 or 33-1-69416704
EMail: roland@lefebvre.ppm.u-psud.fr

VOLIA LEMOS
Universidade Estadual de Campinas
Instituto de Fisica 'Glen Wataguin'
CeP 13083-970
Campinas, SP
Brazil
Phone: 55-19-2393137
Fax: 55-19-2393137
EMail: volia@ifi.unicamp.br

JERZY LESZCZYNSKI
Jackson State University
Dept. of Chemistry
1400 Lynch Street
Jackson, MS 39217
USA
Phone: 601-973-3482
Fax: 601-973-3674
EMail: jerzy@tiger.jsums.edu

ANDREW LIASHENKO
Florida State University
SCRI
Box 152
Tallahassee, FL 32306
USA
Phone: 904-644-2782
Fax: 904-644-0068
EMail: andrew@scri.fsu.edu

ANTONIO TADEU LINO
Universidad Federal de Uberlandia
Depto. de Ciencias Fisicas
Campus Santa Monica
Uberlandia, MG 38400-902
Brazil
Phone: 55-34-235-2888
Fax: 55-34-239-4106
EMail: atlino@inga.ufu.br

IGOR LITVINYUK
Florida State University
Chemistry
Institute of Molecular Biophysics
Tallahassee, FL 32306-3015
USA
Phone: 904-644-6452
Fax: 904-644-3257

GUANGHUA LIU
Florida State University
SCRI
495 Dirac Science Library
Tallahassee, FL 32306-4052
USA
Phone: 904-644-7060
Fax: 904-644-0098
EMail: gliu@scri.fsu.edu

GILDA LOEW
Molecular Research Institute
845 Page Mill Road
Palo Alto, CA 94304
USA
Phone: 415-424-9924
Fax: 415-424-9501
EMail: loew@montara.molres.org

PER-OLOV LÖWDIN
University of Florida
Quantum Theory Project
365 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: lowdin@qtp.ufl.edu

HECTOR LUNA-GARCÍA
Universidad Autonoma Metropolitana-Azcapotzalco
Ciencias Basicas
Av. San Pablo 180
Mexico, DF 02200
Mexico
Phone: 915-724-4218
EMail: lghm@hp9000al.uam.mx

GERALD LUSHINGTON
Army Research Laboratory
AMSRL-WM-PC
Aberdeen, MD 21005-5066
USA
Phone: 410-278-6182
Fax: 410-278-6150
EMail: ghl@arl.mil

UKO MARAN
University of Florida
Department of Chemistry
229 Chemistry Research Bldg.
Gainesville, FL 32611-7200
USA
Phone: 352-392-0554
Fax: 352-392-9199
EMail: uko@ufark2.chem.ufl.edu

NORMAN MARCH
Oxford University
6 Northcroft Road
Egham, Surrey TW20 ODU
UK

RUDOLPH A. MARCUS
California Institute of Technology
Division of Chemistry
110 Noyes Lab. Chem. Physics
Pasadena, CA 91125
USA
Phone: 818-395-6566
Fax: 818-792-8485
EMail: ram@caltech;edu

JAN MARKUSSEN
Copenhagen University
Department of Chemistry
Universitetsparken 5
Copenhagen DK-2100
Denmark
Phone: 45-3-532-0286
EMail: jan@kl4aix.ki.ku.dk

JOSE MARTÍNEZ-MAGADÁN
Instituto Mexicano del Petroleo
SIGA, Eje Central Lazaro Cardenas 152, AP 14805
Mexico, DF 07730
Mexico
Phone: 52-3685911

LOU MASSA
City University of New York
Dept. of Chemistry
695 Park Avenue
New York, NY 10021
USA
Phone: 212-772-5330
Fax: 212-772-5332
EMail: massa@mvaxgr.hunter.cuny.edu

RICHARD J. MATHAR
University of Florida
Quantum Theory Project
368 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: mathar@qtp.ufl.edu

FRANCISCO JAVIER MELÉNDEZ BUSTAMANTE
Inst. de Estructura de la Materia CSIC
Lab. de Quimica Cuantica
Serrano 123
Madrid E-28006
Spain
Phone: 34-1-585-54-04
Fax: 34-1-585-51-84
EMail: emmelendez@roca.csic.es

CELSO PINTO DE MELO
Universidade Federal de Pernambuco
Departamento de Fisica
50 670-901 Recife PE Brazil
Recife, Pernambuco
Brazil
Phone: 1-55-81-271-8450
Fax: 1-55-81-271-0359
EMail: celso@df.ufpe.br

DAVID A. MICHA
University of Florida
Quantum Theory Project
366 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-2597
Fax: 352-392-8722
EMail: micha@qtp.ufl.edu

HENDRIK J. MONKHORST
University of Florida
Quantum Theory Project
362 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: monkhors@qtp.ufl.edu

MARCO ANTONIO MORA-DELGADO
Universidad Autonoma Metropolitana - Iztapalapa
Dpto. de Quimica
Av. Michoacan y La Purisima, Col Vicentina
Iztapalapa, D F C P 09340
Mexico
Phone: 52-5-724-4675
Fax: 52-5-724-4666
EMail: mam@xanum.uam.mx

JORGE A. MORALES
University of Florida
Quantum Theory Project
348 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: morales@qtp.ufl.edu

JOHN D. MORGAN
University of Delaware
Dept. of Physics & Astronomy
Newark, DE 19716
USA
Phone: 302-831-2661
Fax: 302-831-1637
EMail: 32399@udel.edu

AKIHIRO MORITA
Kyoto University
Dept. of Chemistry
Kitashirakawa Sakyo-Ku
Kyoto 606
Japan
Phone: 81-75-753-4006
EMail: morita@kuchem.kyoto-u.ac.jp

RONALDO MOTA
Universidade Federal de Santa Maria
Dept. de Fisica
Santa Maria, RS 97119-900
Brazil
Phone: 55-55-2261616
Fax: 55-55-2262919
EMail: mota@super.ufsm.br

BENGT NAGEL
Royal Institute of Technology
Dept. of Physics
Stockholm S-100 44
Sweden
Phone: 46-8-7907168
Fax: 46-8-104879
EMail: nagel@theophys.kth.se

DENNIS M. NEWNS
IBM
T. J. Watson Laboratory
P.O. Box 218
Yorktown Heights, NY 10598
USA
Phone: 914-945-3551
Fax: 914-945-2141
EMail: dmnewns@watson.ibm.com

MARSHALL NEWTON
Brookhaven National Laboratory
Dept. of Chemistry
P.O. Box 5000
Upton, NY 11973
USA
Phone: 516-344-4366
Fax: 516-344-5815
EMail: newton@bnl.gov

SONJA NIKOLIC
The Rugjer Boskovic Institute
Physical Chemistry Dept.
Bijenicka 54
Zagreb 10001
Croatia
Phone: 385-1-424689
Fax: 385-1-272648
EMail: sonja@olimp.irb.hr

MARCEL NOOIJEN
University of Florida
Quantum Theory Project
343 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: nooijen@qtp.ufl.edu

TED O'BRIEN
University of Florida
Quantum Theory Project
201 CRB, P.O. Box 118435
Gainesville, FL 32611
USA
Phone: 352-392-9306
Fax: 352-392-8722
EMail: obrien@qtp.ufl.edu

YNGVE ÖHRN
University of Florida
Quantum Theory Project
363 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: ohrn@qtp.ufl.edu

KOJI OHTA
Osaka National Research Institute, AIST, MITI
Dept. of Optical Materials
1-8-31 Midorigaoka
Ikeda, Osaka 563
Japan
Phone: 81-6-727-51-9523
Fax: 81-6-727-51-9628
EMail: ohta@onri.go.jp

JOSEP M. OLIVA
University of Bristol
Dept. of Chemistry
Cantocks Close
Bristol BS8 1TS
UK
Phone: 44-117-928-9000 EXT 4271
Fax: 44-117-925-1295
EMail: j.m.oliva@bristol.ac.uk

MONICA OLIVA
Universitat Jaune I
Dept. de Ciencies Experimentals
Box 242
Castellon 12080
Spain
Phone: 34-64-345741
Fax: 34-64-345654
EMail: oliva@vents.uji.es

JOSE ONUCHIC
University of California, San Diego
Dept. of Physics
9500 Gilman Drive
La Jolla, CA 92093
USA
Phone: 619-534-7067
Fax: 619-534-7697
EMail: jonuchic@ucsd.edu

VINCENT ORTIZ
Kansas State University
Dept. of Chemistry
Manhattan, KS 66506-3701
USA
Phone: 913-532-6665
Fax: 913-532-6666
EMail: ortiz@ksu.edu

ROMAN OSMAN
Mt. Sinai School of Medicine
Dept. Physiology & Biophysics
Box 1218
New York, NY 10029
USA
Phone: 212-241-5609
Fax: 212-860-3369
EMail: osman@msvax.mssm.edu

NEIL OSTLUND
Hypercube, Inc.
419 Phillip St., Unit 7
Waterloo, Ontario N2L 3X2
Canada
Phone: 519-725-4040

GEORGE R. PACK
University of Illinois
College of Medicine
1601 Parkview Ave.
Rockford, IL 61107
USA
Phone: 815-395-5694
Fax: 815-395-5666
EMail: georgep@uic.edu

JOSEPH M. PAIKEDAY
Southeast Missouri State University
Dept. of Physics
One University Plaza MS 6600
Cape Girardeau, MO 63701-4799
USA
Phone: 573-651-2393
Fax: 573-651-2223
EMail: c314scp@semovm.semo.edu

GREG MARTIN PEARL
University of Florida
Quantum Theory Project
P.O. Box 118435
Gainesville, FL 32611
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: pearl@qtp.ufl.edu

AJITH S. PERERA
University of Florida
Quantum Theory Project
343 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: perera@qtp.ufl.edu

WILLIS PERSON
University of Florida
Department of Chemistry
P.O. Box 117200
Gainesville, FL 32611-7200
USA
Phone: 352-392-0528

PAULO PIQUINI
Universidade Federal de Santa Maria
Dpto. de Fisica
Santa Maria, RS 97119-900
Brazil
Phone: 55-55-2262919
Fax: 55-55-2262919
EMail: ppiquini@oslo.ccne.ufsm.br

PAWEL PISKORZ
Florida State University
SCRI
495 Dirac Science Library
Tallahassee, FL 32306-4052
USA
Phone: 904-644-4060
Fax: 904-644-0098
EMail: pawel@scri.fsu.edu

EMIL POP
Pharmos Corporation
2 Innovation Drive
Alachua, FL 32615
USA
Phone: 352-462-1210
Fax: 352-462-5401
EMail: emilpop@aol.com

JOHN POPLE
Northwestern University
Dept. of Chemistry
2145 Sheridan Road
Evanston, IL 60208-3113
USA
Phone: 847-491-3403
Fax: 847-491-7713
EMail: pople@lithium.chem.nwu.edu

LEONARD E. PORTER
Washington State University
Radiation Safety Office
Pullman, WA 99164-1302
USA
Phone: 509-335-7057
Fax: 509-335-1615
EMail: porterl@mail.wsu.edu

ENRIQUE POULAIN
Instituto Technologico de Tlalnepantla
Division de Estudios de Posgrado
Apdo Postal 750
Tlalnepantla de Baz 54070
Mexico
Phone: 52-5-390-0310
Fax: 52-5-565-3910
EMail: sca@hp9000a1.uam.mx

MICHAEL PROBST
Innsbruck University
Dept. of Inorganic Chemistry
Innrain 520
Innsbruck A-6020
Austria
Phone: 43-512-5075153
Fax: 43-512-5072934
EMail: michael.probst@uibk.ac.at

PETER PULAY
University of Arkansas
Department of Chemistry
Fayetteville, AR 72701-1201
USA
Phone: 501-575-6612
EMail: pp24139@uafsysb.uark.edu

GEORGE PURVIS
Oxford Molecular
P.O. Box 4003
Beaverton, OR 97076
USA
Phone: 503-526-5000
Fax: 503-526-5099
EMail: gpurvis@oxmol.com

CARLOS QUINTANAR
Universidad Nacional Autonoma de México
Facultad de Ciencias
Mexico, DF C.P.04510
Mexico
Phone: 525-622-50-86
Fax: 525-616-15-35
EMail: cqs@moon.iflsicacu.unam.mx

JAMES RABINOWITZ
U.S. Environmental Protection Agency
MD-68, Health Effects Research Laboratory
Research Triangle Park, NC 27711
USA
Phone: 919-541-5714
Fax: 919-541-0694
EMail: sar@linus.herl.epa.gov

MICHAEL RAMEK
Technische Universität Graz
Phys. & Theor. Chemistry
Graz A-8010
Austria
Phone: 43-316-873-8227
Fax: 43-316-873-8720
EMail: ramek@ptc.tu-graz.ac.at

MILAN RANDIC
Drake University
Dept. of Math & Computer Science
Des Moines, IA 50311
USA
Phone: 515-271-2163
Fax: 515-271-2055

VITALY RASSOLOV
Northwestern University
Dept. of Chemistry
2145 Sheridan Road
Evanston, IL 60208
USA
Phone: 847-491-3423
Fax: 847-491-7713
EMail: rassolov@chem.nwu.edu

JOSE RÉCAMIER
Universidad Nacional Autonoma de Mexico
Lab Cuernavaca
Apdo. Postal 48-3
Cuernavaca, Morelos 62251
Mexico
Phone: 52-5-622-7763
Fax: 52-7-317-3077
EMail: pepe@ce.itisicam.unam.mx

WILLIAM RHODES
Florida State University
Dept. of Chemistry
Tallahassee, FL 32306-3006
USA
Phone: 904-644-1227
Fax: 904-644-8281
EMail: rhodes@chem.fsu.edu

RANDY RICH
Silicon Graphics
2011 N. Shoreline Blvd.
Mountainview, CA 94039-7311
USA
Phone: 415-933-1772
Fax: 415-932-1772

NIGEL RICHARDS
University of Florida
Dept. of Chemistry
P.O. Box 117200
Gainesville, FL 32611-7200
USA
Phone: 352-392-3601
Fax: 352-846-2095
EMail: richards@qtp.ufl.edu

UWE RISS
University of Heidelberg
Physikalisch-Chemisches Institut
Heidelberg 69120
Germany
Phone: 49-6221-545209
Fax: 49-6221-545209
EMail: uwer@tc.pci.uni-heidelberg.de

JUVENCIO ROBLES
University of Guanajuato
Facultad de Quimica
Noria Alta A/N
Guanajuato, GTO 36050
Mexico
Phone: 52-473-26885
Fax: 52-473-24250
EMail: roblesj@quijote.ugto.mx

NOTKER RÖSCH
Technische Universität München
Inst. fur Physik. & Theo. Chem.,
Lichtenbergstrasse 4
Garching D-85747
Germany
Phone: 49-89-2891-3616
Fax: 49-89-2891-3622
EMail: rosch@qtp.ufl.edu

RICHARD B. ROSS
PPG Industries Inc.
4325 Rosanna Drive, P.O. Box 9
Allison Park, PA 15101
USA
Phone: 412-492-5359

PIOTR B. ROZYCZKO
University of Florida
Quantum Theory Project
345 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: rozyczko@qtp.ufl.edu

KEITH RUNGE
University of Florida
Quantum Theory Project
376 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: runge@qtp.ufl.edu

JOHN R. SABIN
University of Florida
Quantum Theory Project
355 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: sabin@qtp.ufl.edu

RAYMOND SADEGHI
University of Florida
Quantum Theory Project
379 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: rsadeghi@qtp.ufl.edu

VIRAHT SAHNI
Brooklyn College of CUNY
Department of Physics
2900 Bedford Avenue
Brooklyn, NY 11210-2889
USA
Phone: 718-951-5785
Fax: 718-951-4407
EMail: vvvbc@cunyvm.cuny.edu

SHOGO SAKAI
Osaka Sangyo University
Dept. of Information Systems Engineering
Daito 574
Japan
Phone: 81-720-89-2538
Fax: 81-720-89-2539
EMail: sakai@ise.osaka-sandai.ac.jp

JULIO R. SAMBRANO
Universidade Estadual Paulista
Depto de Matematica
CP 473
Bauru, SP 17030 360
Brazil
Phone: 55-16-2721561
EMail: pjrs@power.ufscar.br

RICHARD SAYKALLY
University of California, Berkeley
Dept. of Chemistry
419 Latimer Hall MC 1460
Berkeley, CA 94720-1460
USA
Phone: 510-642-8269
Fax: 510-642-8369
EMail: saykally@cchem.berkeley.edu

HAROLD SCHERAGA
Cornell University
Baker Lab. of Chemistry
Ithaca, NY 14853-1307
USA
Phone: 607-255-4034
Fax: 607-254-4700
EMail: has5@cornell.edu

PETER SCHMIDT
Office of Naval Research
Chemistry Division
Code 1113, 800 North Quincy Street
Arlington, VA 22217-5000
EMail: schmidt@onrhq.onr.navy.mil

WILLIAM F. SCHNEIDER
Ford Motor Co.
Chemistry Dept.
P.O. Box 2053 MD 3083/SLR
Dearborn, MI 48121-2053
USA
Phone: 313-323-2064
Fax: 313-594-2923
EMail: wschnei2@ford.com

DIETER SCHUCH
J.W. Goethe-Universität
Inst. fur Theoretische Physik
Robert-Mayer
D-60054 Frankfurt Am Main
Germany
Phone: 49-69-319523
Fax: 49-69-3088997

KLAUS SCHULTEN
University of Illinois
Beckman Institute
405 N. Mathews Avenue
Urbana, IL 61801
USA
Phone: 217-244-1604
Fax: 217-244-6078
EMail: kschulte@ks.uiuc.edu

SANJA SEKUSAK
Rugjer Boskovic Institute
Dept. of Chemistry
Bijenicka 54, P.O. Box 10
Zagreb
Croatia
Phone: 385-1-4561-089
Fax: 385-1-272-648
EMail: sanja@indigo.irb.hr

JORGE M. SEMINARIO
University of New Orleans
Department of Chemistry
New Orleans, LA 70148
USA
Phone: 504-286-7216
Fax: 504-286-6860
EMail: jsmcm@uno.edu

EUGENE SHAKHNOVICH
Harvard University
Depts. of Chemistry & Chemical Biology
12 Oxford Street
Cambridge, MA 02138
USA
Phone: 617-495-4130
Fax: 617-496-5948
EMail: eugene@diamond.harvard.edu

YASU TERU SHIGETA
Kanazawa University
Dept. of Chemistry
Kakuma Machi
Kanazawa 920-11
Japan
Phone: 81-732-64-5686
EMail: kiyoshi@wriron1.s.kanazawa-u.ac.jp

JEFFREY SKOLNICK
Scripps Research Institute
Dept. of Molecular Biology
10550 North Torrey Pines Road
La Jolla, CA 92037
USA
Phone: 619-784-8821
Fax: 619-784-8895
EMail: skolnick@scripps.edu

YVES G. SMEYERS
Superior Council for Scientific Investigations
Institute of Matter Structure
Calle Serrano No. 123
Madrid E-28006
Spain
Phone: 34-1-5855404
Fax: 34-1-5642431
EMail: emsmeyers@roca.csic.es

JOSE M. SOLER
Universidad Autonoma de Madrid
Dept. Fisica Mat. Cond.
Madrid E-28049
Spain
Phone: 34-1-3975550
Fax: 34-1-3973961
EMail: jose@polar.fmc.uam.es

RAMON SOSA
Universidad de la Republica
Inst. of Physics
Ada Herrera Reissig 565
Montevideo 11300
Uruguay
Phone: 598-2-710905
Fax: 598-2-711630
EMail: rsosa@fing.edu.uy

RICHARD SQUIRE
Marshall University
Chemistry Department
901 W. DuPont Avenue
Belle, WV 25015
USA
Phone: 304-965-3084
Fax: 304-357-1230

VOLKER STAEMMLER
Ruhr-Universität Bochum
Lehrstuhl fur Theoretische Chemie
Universitatsstr 150
Bochum D-44801
Germany
Phone: 49-234-7006752
Fax: 49-234-7094109
EMail: staemm@mokat50.theochem.ruhr-uni-bochum.de

KRASSIMIR K. STAVREV
University of Florida
Quantum Theory Project
380 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: stavrev@qtp.ufl.edu

WALTER STEVENS
National Institute of Standards and Technology
Building 221, Room A111
Gaithersburg, MD 20899
USA
Phone: 301-975-5968
Fax: 301-869-4020
EMail: walter.stevens@nist.gov

HEGE STROMSNAS
University of Florida
Quantum Theory Project
P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722

ALEXEI A. STUCHEBRUKHOV
University of California, Davis
Dept. of Chemistry
Davis, CA 95616
USA
Phone: 916-752-7778
Fax: 916-752-8995
EMail: stuchebr@chem.ucdavis.edu

TZU-MIN SU
National Taiwan University
Dept. of Chemistry
Taipei
Taiwan
Phone: 886-2-3630231
Fax: 886-2-3636359
EMail: tmsu@rs350.ch.ntu.edu.tw

JUN-QIANG SUN
University of Florida
Quantum Theory Project
386 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: sun@qtp.ufl.edu

KRYSTYNA SZCZEPANIAK-PERSON
University of Florida
Dept. of Chemistry
P.O. Box 117200
Gainesville, FL 32611-7200
USA
Phone: 352-392-4653
Fax: 352-392-0872
EMail: ksperson@pine.circa.ufl.edu

JAMES TALMAN
University of Western Ontario
Dept. of Applied Mathematics
WSC 173
London, Ontario N6A 5B7
Canada
Phone: 519-679-2111, ext. 8800
Fax: 519-661-3523
EMail: jdt@apmaths.uwo.ca

BRIAN THORNDYKE
University of Florida
Computer and Inf. Sciences
301 CSE
Gainesville, FL 32611
USA
Phone: 352-392-1435
EMail: thorndyb@cise.ufl.edu

KATARZYIVA TKACZ-SMIECH
Technical University of Mining and Metallurgy
Faculty of Materials Science & Ceramic
AL MICV IEWICZA 30
Krakow 30-059
Poland
Phone: 48-12-173788
Fax: 48-12-331593
EMail: smiech@uci.agh.edu.pl

SAMUEL B. TRICKEY
University of Florida
Quantum Theory Project
364 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: trickey@qtp.ufl.edu

NENAD TRINAJSTIC
Rugjer Boskovic Institute
Dept. of Physics and Chemistry
P.O. Box 1016
Zagreb HR 41001
Croatia
Phone: 385-1-424683
Fax: 385-1-272648
EMail: trina@olimp.irb.hr

JOHN TULLY
Yale University
Dept. of Chemistry
225 Prospect Street
New Haven, CT 06520
USA
Phone: 203-432-3934
Fax: 203-432-6144
EMail: tully@onsager.chem.yale.edu

PIET VAN DUIJNEN
University of Groningen
Dept. of Chemistry
Bldg. 223
Groningen, AG 9747-AG
Netherlands
Phone: 31-50-3634373
Fax: 31-50-3634290
EMail: ptvd@chem.rug.nl

CAROLYNE M. VAN VLIET
Florida International University
Dept. of Elect. & Comp. Engineering
University Park
Miami, FL 33199
USA
Phone: 305-348-3709
Fax: 305-348-3707
EMail: vanvliet@eng.fiu.edu

LASZLO VON SZENTPALY
University of the West Indies
Chemistry Department
Mona Campus
Kingston
Jamaica
Phone: 809-927-1910
Fax: 809-977-1835
EMail: laszlo@uwimona.edu.jm

DAVID WALES
University of Cambridge
University Chemistry Lab.
Lensfield Road
Cambridge CB2 1EW
UK
Phone: 44-01223-336354
Fax: 44-01223-336362
EMail: dw34@clust.ch.cam.uc.uk

JIAN WANG
University of Florida
Quantum Theory Project
P.O. Box 118435
Gainesville, FL 32611
USA
Phone: 352-392-8113
Fax: 352-392-8722
EMail: wang@qtp.ufl.edu

ARIEH WARSHEL
University of Southern California
Chemistry Department
University Park
Los Angeles, CA 90089
USA
Phone: 213-740-4114

JOHN WATTS
University of Florida
Quantum Theory Project
367 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-6714
Fax: 352-392-8722
EMail: watts@qtp.ufl.edu

CHARLES WEATHERFORD
Florida A & M University
Dept. of Physics
205 Jones Hall
Tallahassee, FL 32307
USA
Phone: 904-599-3470
Fax: 904-599-3577
EMail: weatherf@ccnnas.nhmfl.gov

BRIAN WEINER
Pennsylvania State University
Dept. of Physics
College Place
Dubois, PA 15801
USA
Phone: 814-375-4700
Fax: 814-234-4291
EMail: bqw@psu.edu

HAREL Y. WEINSTEIN
Mt. Sinai School of Medicine
Physiology & Biophysics Dept.
One Gustave L. Levy Place
New York, NY 10029-6574
USA
Phone: 212-241-7018
Fax: 212-860-3369
EMail: hweinstein@msvax.mssm.edu

HORST WEISS
BASF Aktiengesellschaft
Zk/C - Tagungen, Bldg. B1
Ludwigshafen 67056
Germany
Phone: 49-621-60-41949
Fax: 49-621-60-21863
EMail: weiss@steve.zk.basf-ag-de

KENNETH WILSON
University of Florida
Quantum Theory Project
P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: wilson@qtp.ufl.edu

PETER WINKLER
University of Nevada
Department of Physics 220
Reno, NV 89557-0058
USA
Phone: 702-784-6792
Fax: 702-784-1398
EMail: winkler@rigel.physics.unr.edu

LINDA WONG
University of Illinois
College of Medicine
1601 Parkview Avenue
Rockford, IL 61107
USA
Phone: 815-395-5691
Fax: 815-395-5666
EMail: lindaw@uic.edu

MARGARET WORKMAN
Purdue University
Department of Chemistry
1393 Brown Building
West Lafayette, IN 47906
USA
Phone: 317-496-2409
Fax: 317-494-0239
EMail: workman@coriolis.atms.purdue.edu

DAVID YARON
Carnegie-Mellon University
Department of Chemistry
4400 5th Avenue
Pittsburgh, PA 15213
USA
Phone: 412-268-1351
Fax: 412-268-1061
EMail: yaron@andrew.cmu.edu

ANTHONY YAU
University of Florida
Quantum Theory Project
P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-7184
Fax: 352-392-8722
EMail: yau@qtp.ufl.edu

TAI-SHENG YEH
National Taiwan University
Dept. of Chemistry
Taipei 106
Taiwan
Phone: 886-2-3630231 EXT 2326
Fax: 886-2-3636359
EMail: tsyeh@rs350.ch.ntu.edu.tw

ZHIGANG YI
University of Florida
Quantum Theory Project
348 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: yi@qtp.ufl.edu

YASUNORI YOSHIOKA
Osaka University
Dept. of Chemistry
Toyonaka, Osaka 560
Japan
Phone: 81-6-850-5807
Fax: 81-6-843-7744
EMail: yyoshi@chem.sci.osaka-u.ac.jp

ANGELICA G. ZACARIAS
Universidad Nacional Automona de Mexico
Dept. de Fisica y Quimica Teorica
Mexico, DF CP 04510
Mexico
Phone: 52-5-622-3783
Fax: 52-5-616-2010
EMail: zacarias@papalotl.pquim.unam.mx

KARL ZACHARY
University of Florida
Dept. of Chemistry
P.O. Box 117200
Gainesville, FL 32611-7200
USA
Phone: 352-392-5876
Fax: 352-392-0872
EMail: karl@physical4.chem.ufl.edu

MICHAEL C. ZERNER
University of Florida
Quantum Theory Project
382 Williamson Hall, P.O. Box 118435
Gainesville, FL 32611-8435
USA
Phone: 352-392-1597
Fax: 352-392-8722
EMail: zerner@qtp.ufl.edu

JOHN ZHANG
New York University
Dept. Chemistry
4 Washington Place, Room 514
New York, NY 10003
USA
Phone: 212-998-8412
Fax: 212-260-7905
EMail: zhang@risc.nyu.edu

XUEHE ZHENG
University of Florida
Quantum Theory Project
201 CRB, P.O. Box 118435
Gainesville, FL 32611
USA
Phone: 352-392-9306
Fax: 352-392-8722
EMail: zheng@qtp.ufl.edu

ALEX ZUNGER
National Renewable Energy Laboratory
1617 Cole Blvd.
Golden, CO 80401
USA
Phone: 303-384-6531
EMail: alex-zunger@nrel.gov

WOJCIECH H. ZUREK
Los Alamos National Laboratory
P.O. Box 1663
Los Alamos, NM 87545
USA
Phone: 505-667-6837
Fax: 505-665-3003
EMail: whz@lanl.gov

International Journal of

Quantum Chemistry

Volume 65
1997

Editor-in-Chief

Per-Olov Löwdin
University of Florida at Gainesville, USA
Uppsala University, Sweden

Editors

Erkki Brändas
Uppsala University, Sweden
Yngve Öhrn
University of Florida at Gainesville, USA

Associate Editors

Osvaldo Goscinski
Uppsala University, Sweden
Sten Lunell
Uppsala University, Sweden
John R. Sabin
University of Florida at Gainesville, USA
Michael C. Zerner
University of Florida at Gainesville, USA

Honorary Editors

Gerhard Herzberg
National Research Council, Ottawa, Ontario, Canada
Kenichi Fukui
Institute for Fundamental Chemistry at Kyoto, Japan
Jerome Karle
Naval Research Laboratory at Washington, DC, USA
Rudy Marcus
California Institute of Technology at Pasadena, USA

Editorial Board

Jiri Čížek
University of Waterloo, Ontario, Canada
Enrico Clementi
Université Louis Pasteur, Strasbourg, France
Raymond Daudel
Académie Européene de Arts, des Sciences et des Lettres, Paris, France
Ernest Davidson
Indiana University at Bloomington, USA
George G. Hall
University of Nottingham, UK
Laurens Jansen
Kusnacht, Switzerland
Norman H. March
University of Oxford, UK
Roy McWeeny
Università di Pisa, Italy
Saburo Nagakura
Graduate University for Advanced Studies, Yokohama, Japan
Kimio Ohno
Hokkaido Information University, Japan
Josef Paldus
University of Waterloo, Ontario, Canada
Robert G. Parr
University of North Carolina at Chapel Hill, USA
Ruben Pauncz
Technion, Haifa, Israel
John A. Pople
Northwestern University at Evanston, Illinois, USA
Alberte Pullman
Institut de Biologie Physico-Chimique, Paris, France
Paul von Ragué Schleyer
Universität Erlangen-Nürnberg, Erlangen, Germany
Harrison Shull
Naval Postgraduate School, Monterey, California, USA
Tang Au-Chin
Jilin University, Changchun, China
Rudolf Zahradník
Czech Academy of Sciences, Prague, Czech Republic

Advisory Editorial Board

Teijo Åberg
Helsinki University of Technology, Espoo, Finland
Michail V. Basilevsky
Karpov Institute of Physical Chemistry, Moscow, Russia
Axel D. Becke
Queen's University, Kingston, Ontario, Canada
Gian Luigi Bendazzoli
Università di Bologna, Italy
Geza Biczo
Hungarian Academy of Sciences, Budapest, Hungary
Jerzy Cioslowski
The Florida State University at Tallahassee, USA
Timothy Clark
Universität Erlangen-Nürnberg, Germany
Giorgina Corongiu
Centro di Richerche, Sviluppo e Studi Superiori in Sardegna, Cagliari, Italy
Mireille Defranceschi
DPEL/SERGD/LMVT, Fontenay Aux Roses, France
Karl F. Freed
The University of Chicago, Illinois, USA
Peter Fulde
Max-Planck-Institut für Physik Komplexer Systeme, Dresden, Germany
Odd Gropen
University of Tromsø, Norway
Trygve Helgaker
University of Oslo, Norway
Ming-Bao Huang
Academia Sinica, Beijing, People's Republic of China
James T. Hynes
University of Colorado at Boulder, USA
Mu Shik Jhon
Korea Advanced Institute of Science and Technology, Seoul, Korea
Hiroshi Kashiwagi
Kyushu Institute of Technology, Fukuoka, Japan
Peter A. Kollman
University of California at San Francisco, USA
Eugene S. Kryachko
Academy of Sciences of Ukraine, Kiev, Ukraine
Sven Larsson
Chalmers University of Technology, Gothenburg, Sweden
Lucas Lathouwers
Universitair Centrum (RUCA), Antwerp, Belgium
Shyi-Long Lee
National Chung Chang University, Taiwan, Republic of China
Claude Leforestier
Université Paris-Sud, Orsay, France
Josef Michl
University of Colorado at Boulder, USA
Nimrod Moiseyev
Israel Institute of Technology, Israel
John D. Morgan III
University of Delaware at Newark, USA
Cleanthes A. Nicolaides
National Hellenic Research Foundation, Greece
J. Vincent Ortiz
Kansas State University at Manhattan, USA
Lars Pettersson
University of Stockholm, Sweden
Leon Phillips
University of Canterbury, Christchurch, New Zealand
Martin Quack
ETH Zürich, Switzerland
Leo Radom
Australian National University, Australia
William Reinhardt
University of Washington at Seattle, USA
Sten Rettrup
H. C. Ørsted Institut, Copenhagen, Denmark
C. Magnus L. Rittby
Texas Christian University at Fort Worth, USA
Michael Robb
King's College, London, UK
Mary Beth Ruskai
University of Massachusetts at Lowell, USA
Harold Scheraga
Cornell University at Ithaca, New York, USA
Vipin Srivastava
University of Hyderabad, India
Nicolai F. Stepanov
Moscow State University, Russia
Jiazong Sun
Jilin University, Changchun, People's Republic of China
Colin Thomson
University of St. Andrews, Scotland, UK
Donald G. Truhlar
University of Minnesota at Minneapolis, USA
Frank Weinhold
University of Wisconsin at Madison, USA
Harel Weinstein
Mount Sinai School of Medicine, New York, USA
Peter Wolynes
University of Illinois at Urbana, USA
Robert E. Wyatt
The University of Texas at Austin, USA

Editorial Production, John Wiley & Sons: Sheila Kaminsky

The *International Journal of Quantum Chemistry* (ISSN 0020-7608) is published semi-monthly with one extra issue in January, March, May, July, August, and November by John Wiley & Sons, Inc., 605 Third Avenue, New York, New York 10158.

Copyright © 1997 John Wiley & Sons, Inc. All rights reserved. No part of this publication may be reproduced in any form or by any means, except as permitted under section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the publisher, or authorization through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (508) 750-8400, fax (508) 750-4470. Periodicals postage paid at New York, NY, and at additional mailing offices.

The code and the copyright notice appearing at the bottom of the first page of an article in this journal indicate the copyright owner's consent that copies of the article may be made for personal or internal use, or for the personal or internal use of specific clients, on the condition that the copier pay for copying beyond that permitted by Sections 107 or 108 of the US Copyright Law.

This consent does not extend to the other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective work, or for resale. Such permission requests and other permission inquiries should be addressed to the Permissions Dept.

Subscription price (Volumes 61–65, 1997): $3,860.00 in the US, $4,160.00 in Canada and Mexico, $4,355.00 outside North America. All subscriptions outside US will be sent by air. Personal rate (available only if there is an institutional subscription): $190.00 in North America, $370.00 outside North America. Subscriptions at the personal rate are available only to individuals. Payment must be made in US dollars drawn on a US bank. Claims for undelivered copies will be accepted only after the following issue has been received. Please enclose a copy of the mailing label. Missing copies will be supplied when losses have been sustained in transit and where reserve stock permits.

Please allow four weeks for processing a change of address. For subscription inquiries, please call (212) 850-6645; e-mail: SUBINFO@Wiley.com.

Postmaster: Send address changes to *International Journal of Quantum Chemistry*, Susan Swanson, Director, Subscription Fulfillment and Distribution, Subscription Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158.

Advertising Sales: Inquiries concerning advertising should be forwarded to Susan Levey, Advertising Sales, John Wiley & Sons, Inc., 605 Third Avenue. New York, NY 10158; (212) 850-8832. Advertising Sales, European Contact: Michael Levermore, Advertising Manager, John Wiley & Sons, Ltd., Baffins Lane, Chichester, Sussex PO 19 1UD, England.

Reprints: Reprint sales and Inquiries should be directed to the customer service department, John Wiley & Sons, Inc. 605 Third Ave., New York, NY 10158. Tel: 212-850-8776.

Manuscripts should be submitted in triplicate and accompanied by an executed Copyright Transfer Form to the Editorial Office, *International Journal of Quantum Chemistry*, Quantum Chemistry Group, Uppsala University, Box 518, S-75120, Uppsala, Sweden. Authors may also submit manuscripts to the Editorial Office, *International Journal of Quantum Chemistry*, Quantum Theory Project, Williamson Hall, University of Florida, Gainesville, Florida 32611. **Information for Contributors** appears in the first and last issue of each volume. **All other correspondence** should be addressed to the *International Journal of Quantum Chemistry*, Publisher, Interscience Division, Professional, Reference, and Trade Group, John Wiley & Sons, Inc., 605 Third Avenue, New York, NewYork 10158, U.S.A. The contents of this journal are indexed or abstracted in *Chemical Abstracts, Chemical Titles, Chemical Database, Current Contents/Physical, Chemical, and Earth Sciences, Research Alert (ISI), Science Citation Index (ISI),* and *SCISEARCH Database (ISI).*

This paper meets the requirements of ANSI/NISO Z39. 48-1992 (Permanence of Paper). ♾

International Journal of Quantum Chemistry Author Index—Volume 65

International Journal of Quantum Chemistry Subject Index—Volume 65

International Journal of

Quantum Chemistry

Contents for Volume 65, 1997

Issue No. 1, 1997

Theoretical and Computational Developments

Issue No. 2, 1997

Theoretical and Computational Developments

Properties, Dynamics, and Electronic Structure of Atoms and Molecules

Issue No. 5, 1997

Issue No. 6, 1997

Published Symposia of the *International Journal of Quantum Chemistry*

1967 QUANTUM CHEMISTRY SYMPOSIUM NO. 1
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory)

1968 QUANTUM CHEMISTRY SYMPOSIUM NO. 2
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory and Quantum Biology)

1969 QUANTUM CHEMISTRY SYMPOSIUM NO. 3 PART 1
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory and Quantum Biology)

1970 QUANTUM CHEMISTRY SYMPOSIUM NO. 3 PART 2
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory and Quantum Biology)

1971 QUANTUM CHEMISTRY SYMPOSIUM NO. 4
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory and Quantum Biology)

1971 QUANTUM CHEMISTRY SYMPOSIUM NO. 5
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory and Quantum Biology)

1972 QUANTUM CHEMISTRY SYMPOSIUM NO. 6
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory and Quantum Biology)

1973 QUANTUM CHEMISTRY SYMPOSIUM NO. 7
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory and Quantum Biology)

1974 QUANTUM CHEMISTRY SYMPOSIUM NO. 8
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory and Quantum Statistics)
QUANTUM BIOLOGY SYMPOSIUM NO. 1
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1975 QUANTUM CHEMISTRY SYMPOSIUM NO. 9
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory and Quantum Statistics)
QUANTUM BIOLOGY SYMPOSIUM NO. 2
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1976 QUANTUM CHEMISTRY SYMPOSIUM NO. 10
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory and Quantum Statistics)
QUANTUM BIOLOGY SYMPOSIUM NO. 3
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1977 QUANTUM CHEMISTRY SYMPOSIUM NO. 11
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory, Collision Phenomena, and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 4
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1978 QUANTUM CHEMISTRY SYMPOSIUM NO. 12
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory, Collision Phenomena and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 5
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1979 QUANTUM CHEMISTRY SYMPOSIUM NO. 13
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory, Collision Phenomena, Quantum Statistics, and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 6
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1980 QUANTUM CHEMISTRY SYMPOSIUM NO. 14
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory, Collision Phenomena, Quantum Statistics, and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 7
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1981 QUANTUM CHEMISTRY SYMPOSIUM NO. 15
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory, Collision Phenomena, and Computational Quantum Chemistry)
QUANTUM BIOLOGY SYMPOSIUM NO. 8
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1982 QUANTUM CHEMISTRY SYMPOSIUM NO. 16
(Proceedings of the International Symposium on Quantum Chemistry, Theory of Condensed Matter, and Propagator Methods in the Quantum Theory of Matter)
QUANTUM BIOLOGY SYMPOSIUM NO. 9
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1983 QUANTUM CHEMISTRY SYMPOSIUM NO. 17
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory, Collision Phenomena and Computational Quantum Chemistry)
QUANTUM BIOLOGY SYMPOSIUM NO. 10
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1984 QUANTUM CHEMISTRY SYMPOSIUM NO. 18
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory, and Computational Quantum Chemistry)
QUANTUM BIOLOGY SYMPOSIUM NO. 11
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1985 QUANTUM CHEMISTRY SYMPOSIUM NO. 19
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory, Scattering Problems, Many Body Phenomena, and Computational Quantum Chemistry)
QUANTUM BIOLOGY SYMPOSIUM NO. 12
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1986 QUANTUM CHEMISTRY SYMPOSIUM NO. 20
(Proceedings of the International Symposium on Atomic, Molecular, and Solid-State Theory, Scattering Problems, Many Body Phenomena, and Computational Quantum Chemistry)

1986 QUANTUM BIOLOGY SYMPOSIUM NO. 13
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1987 QUANTUM CHEMISTRY SYMPOSIUM NO. 21
(Proceedings of the International Symposium on Quantum Chemistry, Solid-State Theory, and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 14
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1988 QUANTUM CHEMISTRY SYMPOSIUM NO. 22
(Proceedings of the International Symposium on Quantum Chemistry, Solid-State Theory, and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 15
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1989 QUANTUM CHEMISTRY SYMPOSIUM NO. 23
(Proceedings of the International Symposium on Quantum Chemistry, Solid-State Theory, and Molecular Dynamics)
QUANTUM BIOLOGY SYMPOSIUM NO. 16
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1990 QUANTUM CHEMISTRY SYMPOSIUM NO. 24
(Proceedings of the International Symposium on Quantum Chemistry, Solid State Physics, and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 17
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1991 QUANTUM CHEMISTRY SYMPOSIUM NO. 25
(Proceedings of the International Symposium on Quantum Chemistry, Solid State Physics, and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 18
(Proceedings of the International Symposium on Quantum Biology and Quantum Pharmacology)

1992 QUANTUM CHEMISTRY SYMPOSIUM NO. 26
(Proceedings of the International Symposium on Atomic, Molecular, and Condensed Matter Theory and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 19
(Proceedings of the International Symposium on the Application of Fundamental Theory to Problems of Biology and Pharmacology)

1993 QUANTUM CHEMISTRY SYMPOSIUM NO. 27
(Proceedings of the International Symposium on Atomic, Molecular, and Condensed Matter Theory and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 20
(Proceedings of the International Symposium on the Application of Fundamental Theory to Problems of Biology and Pharmacology)

1994 QUANTUM CHEMISTRY SYMPOSIUM NO. 28
(Proceedings of the International Symposium on Atomic, Molecular, and Condensed Matter Theory and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 21
(Proceedings of the International Symposium on the Application of Fundamental Theory to Problems of Biology and Pharmacology)

1995 QUANTUM CHEMISTRY SYMPOSIUM NO. 29
(Proceedings of the International Symposium on Atomic, Molecular, and Condensed Matter Theory and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 22
(Proceedings of the International Symposium on the Application of Fundamental Theory to Problems of Biology and Pharmacology)

1996 QUANTUM CHEMISTRY SYMPOSIUM NO. 30
(Proceedings of the International Symposium on Atomic, Molecular, and Condensed Matter Theory and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 23
(Proceedings of the International Symposium on the Application of Fundamental Theory to Problems of Biology and Pharmacology)

1997 QUANTUM CHEMISTRY SYMPOSIUM NO. 31
(Proceedings of the International Symposium on Atomic, Molecular, and Condensed Matter Theory and Computational Methods)
QUANTUM BIOLOGY SYMPOSIUM NO. 24
(Proceedings of the International Symposium on the Application of Fundamental Theory to Problems of Biology and Pharmacology)

All of the above symposia can be individually purchased from the Subscription Department, John Wiley & Sons.

Guidelines for Electronic Submission

The *International Journal of Quantum Chemistry* strongly encourages authors to deliver the revised version of their manuscripts (text, tables and, if possible, illustrations) on diskette.

Text

- Storage medium: 3 ½" (preferred) or 5¼" diskette in IBM MS-DOS, Windows, or Macintosh format.
- Software and format: Microsoft Word 6.0 is preferred, although manuscripts prepared using any other microcomputer wordprocessor are acceptable. Refrain from complex formatting; the Publisher will style your manuscript according to the *International Journal of Quantum Chemistry* design specifications. Do not use desktop publishing software such as Aldus Pagemaker or QuarkXPress. If you prepared your manuscript using one of these programs, export the text to a wordprocessing format. Please make sure your wordprocessing program's "fast save" feature is turned off.
- File names: Submit the text and tables of each manuscript as a single file. Name each file with your last name (up to eight letters). Text files should be given the three-letter extension that identifies the file format. Macintosh users should maintain the MS-DOS "eight dot three" file-naming convention.
- Labels: Label all diskettes with your name, the file name, and the wordprocessing program and version used.
- Paper copy: Accompany all files with a printed paper copy.

Illustrations

- Storage medium: Submit as separate files from text files, on separate diskettes or cartridges. 3½" diskettes, Iomega Zip, and 5¼" 44- or 88-MB SyQuest cartridges can be submitted. At authors' request, cartridges and diskettes will be returned after publication.
- Software and format: The preferred formats are TIFF or EPS with pict or tiff preview, although any format that is in general use that is not application-specific is acceptable.
- Resolution: Journal quality reproduction will require greyscale and color files at resolutions yielding approximately 300ppi. Bitmapped line art should be submitted at resolutions yielding 600–1200 ppi. These resolutions refer to the output size of the file; if you anticipate that your images will be enlarged or reduced, resolutions should be adjusted accordingly.
- File names: Illustration files should be given the 2- or 3-letter extension that identifies the file format used (i.e., TIFF, EPS, RGB, etc.).
- Labels: Label all diskettes and cartridges with your name, the file names, formats, sizes, and compression schemes (if any) used. Hard copy output must accompany all files.

Manuscripts and illustrations not conforming to the style of this journal will be returned to the author for reworking, thus delaying their appearance.

??? Questions ???

Sheila Kaminsky
Email: SKaminsk@Wiley.com
Telephone: 212-850-6540
Fax: 212-850-6052

WILEY

Publishers Since 1807

COPYRIGHT TRANSFER AGREEMENT

Date:

To:

Production/Contribution ID# ______ Publisher/Editorial office use only

Re: Manuscript entitled ______________________________________ (the "Contribution")

for publication in ______________________________________ (the "Journal")

published by John Wiley & Sons, Inc.

Dear Contributor(s):

Thank you for submitting your Contribution for publication. In order to expedite the publishing process and enable the Publisher to disseminate your work to the fullest extent, we need to have the following copyright transfer agreement signed and returned to us as soon as possible. If the Contribution is not published in any edition of the Journal, the Agreement shall be null and void.

A. The Contributor hereby transfers to John Wiley & Sons. Inc. (the " Publisher"), during the full term of copyright, the full and exclusive rights comprised in the copyright in and to the Contribution, including but not limited to the right to publish, republish, transmit, sell, distribute and otherwise use the Contribution and the material contained therein in electronic and print editions of the Journal and all derivative works, therefrom throughout the world, in all languages and in all media of expression now known or later developed, and to license or permit others to do so.

B. Notwithstanding the above, the Contributor retains all proprietary rights other than copyright, such as patent rights, in any process, procedure or article of manufacture described in the Contribution.

C. The publisher grants back to the Contributor the following:
   1. The right to make and distribute copies of all or part of the Contribution for the Contributor's use in teaching.
   2. The right to use, *after publication* in the Journal, all or part of the material from the Contribution in a book by the Contributor, or in a collection of the Contributor's work.
   3. The royalty-free right to make copies of the Contribution for internal distribution within the institution/company that employs the Contributor subject to the provisions of paragraph E below for a work-made-for hire.
   4. The right to use figures and tables from the Contribution, and up to 250 words of text, for any purpose.
   5. The right to make oral presentations of material from the Contribution.

D. In the case of a Contribution prepared under U.S. Government contract or grant, the U.S. Government may reproduce, royalty-free, all or portions of the Contribution and may authorize others to do so, for official U.S. Government purposes only, *if the U.S. Government contract or grant so requires.*

E. If the Contribution was written by the Contributor as a work-made-for-hire in the course of employment, the Contribution is owned by the company/employer which must sign this Agreement in the space provided below. In such case, the company/employer hereby transfers and assigns to the Publisher the full and exclusive rights in the Contribution specified in paragraph A above. In addition to the rights specified as retained in paragraph B above, the Publisher hereby licenses back to such company/employer, its subsidiaries and divisions, the royalty-free right to use and distribute the Contribution internally or for promotional and non-commercial purposes only. Copies so made shall not be available for re-sale but may be included by the company/employer as part of an information package for licensing purposes.

F. The Contributor and the company/employer agree that any and all copies of the Contribution or any part thereof published under the terms of paragraphs C and E will include a notice of copyright in the Publisher's name and a citation to the Journal.

G. The Contributor represents that the Contribution is the Contributor's original work. If the Contribution was prepared jointly, the Contributor agrees to inform the co-Contributors of the terms of this Agreement and to obtain their signature to this Agreement or their written permission to sign on their behalf. The Contribution is submitted only to this Journal and has not been published before. (If excerpts from copyrighted works are included, the Contributor will obtain written permission from the copyright owners and show credit to the sources in the Contribution.) The Contributor also warrants that the Contribution contains no libelous or unlawful statements, does not infringe on the rights of others, or contain material or instructions that might cause harm or injury.

Check one:

______ Contributor-owned work	Contributor's signature	Date
	Typed or printed name and title	
	Co-contributor's signature	Date
	Typed or printed name and title	
______ Work-made-for-hire for Employer	Institution or company (Employer-for-Hire)	Date
______ U.S. Government work	Authorized signature of Employer	

ATTACH ADDITIONAL SIGNATURE PAGE AS NECESSARY

Note to U.S. Government Employees

A Contribution prepared by a U.S. federal government employee as part of his/her official duties is called a "U.S. Government work." and is in the public domain in the United States. In such case, Paragraph A above applies only outside the United States. If the Contribution was not prepared as part of the employee's duties, it is not a U.S. Government work.

If the Contribution was prepared jointly, and any co-contributor is not a U.S. government employee, *the Contribution is not a U.S. Government work.* In such case, the co-contributor who is not a non-U.S. government employee should be delegated in writing by the other co-contributors to sign this Agreement on their behalf.

International Journal of Quantum Chemistry
Information for Contributors

1. Manuscripts should be submitted in triplicate and accompanied by an executed Copyright Transfer Form to the Editorial Office, International Journal of Quantum Chemistry, Quantum Chemistry Group, University of Uppsala, Box 518, S-75120 Uppsala, Sweden. Authors may also submit manuscripts to the Editorial Office, International Journal of Quantum Chemistry, Quantum Theory Project, 362 Williamson Hall, University of Florida, Gainesville, Florida 32611.

   All other correspondence should be addressed to the Publisher, Professional, Reference, & Trade Group, John Wiley & Sons, Inc., 605 Third Ave., New York, NY 10158.

2. It is the preference of the Editors that papers be published in the English language. However, if the author desires that his paper be published in French or German, it is necessary that a particularly complete and comprehensive synopsis be furnished in English.

3. Manuscripts should be submitted in triplicate (one *original*, two carbon copies) typed *doubled spaced* throughout and on one side of each sheet only, on a *heavy* grade paper with margins of at least 2.5 cm on all sides. Copyright: No article can be published unless accompanied by a signed publication agreement, which serves as a transfer of copyright from author to publisher. A publication agreement may be obtained from the editor or the publisher. A copy of the publication agreement appears in most issues of the journal. Only original papers will be accepted and copyright in published papers will be vested in the publisher. It is the author's responsibility to obtain written permission to reproduce material that has appeared in another publication. A copy of that agreement, executed and signed by the author, is now required with each manuscript submission. (If the article is a "work made for hire," the agreement must be signed by the employer.)

4. A short synopsis (maximum length 200 words) is required. The synopsis should be a summary of the entire paper, not the conclusions alone. If the paper is written in French or German, a synopsis in English should also be prepared. The paper should be reasonably subdivided into sections and, if necessary, subsections.

5. A list of five key words or phrases for indexing must accompany each submission.

6. Authors are cautioned to type—wherever possible—all mathematical and chemical symbols, equations, and formulas. If these must be handwritten, please print clearly and leave ample space above and below for printer's marks; please use only ink. All Greek or unusual symbols should be identified in the margin the first time they are used. Please distinguish in the margins of the manuscript between capital and small letters of the alphabet wherever confusion may arise (e.g., k, K, k). Please underline with a wavy line all vector quantities. Use fractional exponents to avoid root signs.

   The nomenclature sponsored by the International Union of Pure and Applied Chemistry is requested for chemical compounds. Unit abbreviations should follow the practices of the American Institute of Physics. Chemical bonds should be correctly placed, and double bonds clearly indicated. Valence is to be indicated by superscript plus and minus signs.

7. The references should be numbered consecutively in the order of their appearance and should be complete, including authors' initials and—for unpublished lectures or symposia—the title of the paper, the date, and the name of the sponsoring society. Please compile references on a separate sheet at the end of the manuscript. Abbreviations of journal titles should conform to the *Bibliographic Guide for Editors & Authors* published by the American Chemical Society.

   References should be limited to literature citations. Explanatory or supplementary material should be treated either as footnotes to text or appendices. Examples:

   [1] D. N. Zubarev, *Nonequilibrium Statistical Thermodynamics* (Consultants Bureau, Plenum, New York, 1974).
   [2] H. Adachi, M. Tsukada, and C. Satoko, J. Phys. Soc. Jpn. **45**, 875 (1978).
   [3] K. Fukui, T. Yonezawa, C. Nagata, H. Katou, A. Imamura, and K. Morokuma, in *Introduction to Quantum Chemistry* (Kagakudojin, Kyoto, 1963), Vol. 1, p. 197.

8. A limited number of color figures that are of critical importance and that significantly enhance the presentation will be considered for publication at the publisher's expense. Color separations or

transparencies (negatives or positives) are optimal. Color slides are preferable to color prints. Any cropping of the color figure should be clearly indicated. Final decision on publication of color figures will be at the discretion of the Editor.

9. Each table should be supplied on a separate sheet (not interspersed with text). Please supply numbers and titles for all tables. All table columns should have an explanatory heading.

10. Please supply legends for all figures and compile these on a separate sheet.

11. Figures should be professionally prepared and submitted in a form suitable for reproduction (camera-ready copy). Computer-generated graphs are acceptable only if they have been printed with a good quality laser printer. Artwork is generally reduced so that the type in the figures is about 2.5 mm high. The maximum final size of figures for this journal is 16 x 21 cm after reduction. Good glossy photographs are required for halftone reproductions. If in doubt about the preparation of illustrations suitable for reproduction, please consult the publisher at the address given in paragraph 1.

12. Senior authors will receive 50 reprints of their articles without charge. Additional reprints can be ordered and purchased by filling out the form enclosed with the proof.

13. The publisher will do everything possible to ensure prompt publication. It will therefore be appreciated if manuscripts and illustrations conform from the outset to the style of the journal. Contributors should use the *Style Manual* of the American Institute of Physics; papers will otherwise have to be returned to the author for revision.

Corrected proofs must be sent back to the publisher within two days to avoid the risk of the author's contribution having to be held over to a later issue.

WILEY PHYSICS

WILEY

ALGEBRAIC METHODS IN MOLECULAR & NUCLEAR STRUCTURE PHYSICS

by Alejandro Frank, Universidad National Autónoma de México and Pieter Van Isacker, University of Surrey, England and Grand Accélérateur National d'Ions Lourds, France

This book provides the basic technical background required in the applications of algebraic models in nuclear physics with an emphasis on group-theoretical techniques. The material is divided into three parts of increasing mathematical complexity. Part I is designed to familiarize the reader with concepts and terminology of group theory as well as with the basic features of boson-fermion systems. Part II is devoted to the study of molecular algebraic models and Part III deals with the interacting boson model of nuclear structure.

Contents:

Part I, Schematic Models: Identical Bosons • Non-identical Bosons • Bosons and Fermions • Supersymmetry and F Spin.

Part II, Molecular Models: Diatomic Molecules • Triatomic Molecules • Bose-Fermi Symmetries and Molecular Electronic Spectra.

Part III, Nuclear Models: The Interacting Boson Model • The Neutron-Proton Interacting Boson Model • The Interacting Boson-Fermion Model • Group Theory and the Algebraic Approach • The Unitary and Orthoganal Algebras • Dragt's Theorem

1994 • 488pp. • 0-471-52640-1 • $64.95

ANGULAR MOMENTUM

An Illustrated Guide to Rotational Symmetries for Physical Systems

by William J. Thompson, University of North Carolina, Chapel Hill

Pedagogically consistent and self-contained, this extensively illustrated guide to rotational symmetries is an excellent learning aid for students of physics, chemistry, and mathematics, as well as a valuable resource for researchers in the physical sciences and applied mathematics. Includes *Mathematica* software that allows you to create vivid images of rotational symmetries.

Contents: The Computer Interface • Symmetry in Physical Systems • Mathematical and Quantal Preliminaries • Rotational Invariance and Angular Momentum • Angular Momentum Eigenstates • Angular Momentum in Quantum Systems • Finite Rotations of Angular Momentum Eigenstates • Combining Two Angular Momentum Eigenstates • Irreducible Spherical Tensors and Spin • Recombining Several Angular Momentum Eigenstates

1994 • 461pp. • 0-471-55264-X • $64.95

PRINCIPLES OF SYMMETRY, DYNAMICS, AND SPECTROSCOPY

by William G. Harter, University of Arkansas, Fayetteville

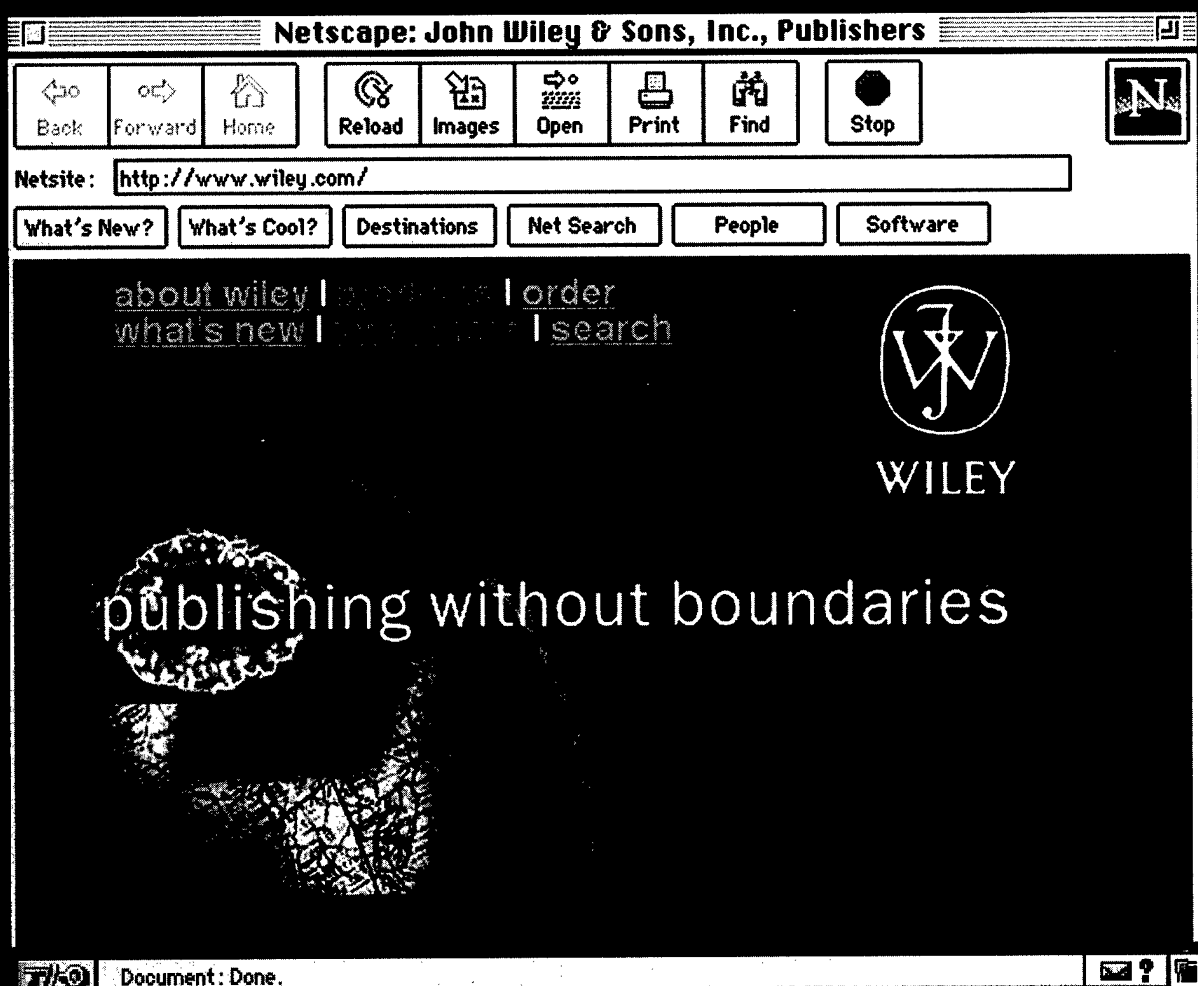
This book is designed to explain the applications of group theory and symmetry analysis to physics by using a novel physical approach. Prof. Harter breaks down the complex mathematics of group theory representation by first introducing the simplest physical problem to show and prove each mathematical concept. These important mathematical quantities are discussed and derived after their need is demonstrated by a physical problem. This approach enables the reader to tackle more complex problems by synthesizing simpler mathematical concepts. This book is a valuable contribution to the rapidly growing field of laser spectroscopy.

Contents: A Review of Matrix Algebra and Quantum Mechanics • Basic Theory and Applications of Symmetry Representations (*Abelian Symmetry Groups*) • Basic Theory and Applications of Symmetry Representations (*Non-Abelian Symmetry Groups*) • Theory and Applications of Higher Finite Symmetry and Induced Representations • Representations of Continuous Rotation Groups and Applications • Theory and Applications of Symmetry Representation Products (*Finite Groups*) • Theory and Application of Symmetry Representation Products (*Continuous Rotation Groups*) • Symmetry Analysis for Semiclassical and Quantum Mechanics: Dynamics with High Quanta

1993 • 846pp. • 0-471-05020-2 • $125.00

JOHN WILEY & SONS, INC.
605 THIRD AVENUE, NEW YORK, NY 10158-0012
To order by phone: (800) 879-4539; To order by fax: (201) 302-2300; Other inquiries (212) 850-6418
Prices subject to change and may be slightly higher outside U.S.A.

JOIN WILEY ONLINE



Drop in and explore. Get the very latest information on new books, journals, and other publications along with special promotions and publicity. Discover a variety of online services, including online journals. Download a sample chapter or software. And much more.

STOP BY TODAY!
www.wiley.com/

WILEY
Publishers Since 1807

0020-7608(1997)65:6;1-0